Structures of Nonequilibrium Fluctuations

Dissipation and Activity

Bram Wynants

Supervisor:
Prof. Dr. C. Maes

Jury:
Prof. Dr. D. Bollé, president
Prof. Dr. T. Bodineau,
(Ecole Normale Supérieure, Paris)
Prof. Dr. M. Fannes
Prof. Dr. J. Indekeu
Prof. Dr. H. Touchette,
(Queen Mary, University of London)
Prof. Dr. W. Troost

U.D.C. number

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Science

2010

D/2010/10.705/26
ISBN 978-90-8649-333-3

Preface

This thesis should in some way summarize the research I have done in these last 4 years as a PhD student. This is done in the next ±200 pages. What those pages do not reveal, is the way in which the results were obtained. For a large part they were obtained by talking to colleagues and discussions with collaborators. These things I get to summarize here, in one page, making an unfair ratio of 1/200.

In the first place I owe many thanks to my supervisor Christian Maes. Christ, thank you for all the time you took to introduce me into this area of research. For being there when I asked for explanation, or discussing possible ideas and results. For stimulating me to go on conferences, giving talks and writing papers. For making these four years a very pleasant and enriching experience.

The fact that I have publications at all is because I had coworkers. Apart from Christ, many thanks to Karel Netočný, who has also been a very pleasant host during my visits to his institute in Prague. Thank you Marco, for providing, apart from useful discussions, simulations to visualize our results: I used your graphs in this thesis. Thanks to Eliran Boksenbojm and Navinder Singh to endure me as a co-author.

I have many pleasant memories of the time I spent with my office-mates throughout these years: Stijn, Jeremy, Eliran and Simi, thank you for all the fun and making me a wizer man even outside of physics.
I would also like to thank the members of the doctoral committee for reading and commenting on this thesis.
The FWO (Fonds Wetenschappelijk Onderzoek Vlaanderen) deserves my thanks for their financial support.

Finally, Mom and Dad: without your support and stimulation I would never have reached this point. To summarize all my gratitude in just two words seems impossible but I will try:

THANK YOU

Abstract

English We discuss research done in two important areas of nonequilibrium statistical mechanics: fluctuation dissipation relations and dynamical fluctuations. The work discussed here was reported before in [2, 3, 4, 70, 71].
First of all, we investigate how expectation values of observables respond to a perturbation, namely a small change in the system's energy. The central object that summarizes this is the response function. In equilibrium systems it is known that this response function can be written in terms of a correlation function in the unperturbed system. This function is a correlation of the observable with the potential that changed the energy of the system. This is called the fluctuation-dissipation theorem. Our contribution here is an investigation of the form of the response function for systems out of equilibrium. We found that the response function can generally be written as the sum of two correlation functions. One correlation function is linked to entropy exchange with the environment, and thus to heat dissipation. The other correlation function has to do with a quantity which we call traffic and which describes in a sense the activity of the system. The results are applied to several explicit examples for which simulations have provided some visualization.
Furthermore, we use the theory of large deviations to examine certain fluctuations in systems out of equilibrium. Large deviation theory provides a natural mathematical framework for equilibrium statistical mechanics. It is therefore useful to investigate that theory also out of equilibrium. We do that by investigating dynamical fluctuations. In dynamical fluctuation theory we consider two kinds of observables: occupations (describing the fraction of time the system spends in each configuration) and currents (describing the changes of configuration the system makes). The probability is then considered that these observables take on some given values different from the typical values. Such a probability often shows an exponential decay in time. The measure for the exponential decay in time, is called the rate function. We explain how to compute the rate functions, and what the physical quantities are that govern their form. As for fluctuation-dissipation relations, entropy and traffic are the main ingredients. Moreover, the rate function that governs the joint probabilities of occupations and currents is

explicitly computed for the classes of models considered and is expressed in terms of entropy and traffic. The rate function for the occupations can be expressed entirely in terms of traffic. We also show that this traffic can be seen as a thermodynamic potential for currents. Finally, for the close-to-equilibrium regime, known variational principles as the minimum entropy production principle are recovered.

Nederlands We bespreken onderzoek in twee belangrijke deelgebieden van de niet-evenwichts statistische mechanica: fluctuatie-dissipatie relaties en dynamische fluctuaties. Dit onderzoek is eerder al besproken in [2, 3, 4, 70, 71].
We onderzoeken hoe verwachtingswaarden van observabelen reageren op een perturbatie, namelijk een kleine verandering van de energie van het systeem. Het centrale object dat dit samenvat is de responsfunctie. In evenwichtssystemen kan deze geschreven worden in termen van een correlatiefunctie in het ongeperturbeerde systeem: een correlatie van de observabele met de potentiaal die de energie van het systeem heeft veranderd. Dit noemt men het fluctuatie-dissipatie theorema. Onze contributie is het onderzoeken van de vorm van de responsfunctie voor systemen uit evenwicht. We vonden dat die geschreven kan worden als de som van twee termen: één is een correlatiefunctie van de observabele met entropie, de ander met traffic, een grootheid die min of meer de activiteit van het systeem beschrijft. De resultaten worden toegepast op enkele expliciete voorbeelden, waarvan simulaties zijn gemaakt ter visualisatie.
We gebruiken de theorie van grote afwijkingen om bepaalde fluctuaties de onderzoeken in systemen uit evenwicht. Voor evenwichts statistische mechanica vormt deze theorie een natuurlijk wiskundig kader. Het is daarom nuttig om ze te ook te onderzoeken voor systemen uit evenwicht. We kijken daarvoor naar dynamische fluctuaties. In de theorie van dynamische fluctuaties beschouwen we twee soorten observabelen: occupaties (die beschrijven hoe lang het systeem in elke configuratie blijft) en stromen (die beschrijven hoe vaak het systeem van configuratie verandert). We beschouwen dan de kans dat deze observabelen bepaalde waarden aannemen verschillend van de typische waarden. Dergelijke kansen vertone vaak exponentieel verval in de tijd. De grootheid die dit exponentieel verval kwantificeert is de fluctuatiefunctionaal. We beschrijven hoe we die fluctuatiefunctionaal kunnen berekenen en wat de fysische grootheden zijn die bepalend zijn voor zijn vorm? Net zoals voor fluctuatie-dissipatie relaties zijn entropie en traffic hier de belangrijke ingredienten. De fluctuatiefunctionaal kan expliciet berekend worden voor beide klassen van modellen die we beschouwen, en kan uitgedrukt worden in termen van entropie en traffic. Wanneer we de fluctuatiefunctionaal berekenen voor de occupaties alleen, zien we dat die volledig uitgedrukt kan worden in termen van traffic. We leggen ook een structuur bloot die sterk doet denken aan de relaties tussen grootheden uit de thermodynamica (thermodynamische potentialen). Tenslotte, voor kleine fluctuaties kunnen we een aantal gekende principes terugvinden: minimum en maximum entropieprincipes.

Nomenclature

k_B	Boltzmann's constant, usually we work in units in which $k_B = 1$,
$\beta = \frac{1}{k_B T}$	inverse temperature,
U, V	potentials,
F, f, g	forces,
W, Q	work and heat,
μ	chemical potential.

Configurations and trajectories

Ω	configuration/state space,
$x, y, \ldots$	configurations/states,
t, s, T	times,
x_t	configuration at time t,
λ_t	protocol,
$\omega = (x_t)_{0 \le t \le T}$	trajectory/path during an interval $[0, T]$,
π	kinematical time-reversal operator changing the signs of velocities,
$\theta\omega = (\pi x_{T-t})_{0 \le t \le T}$	time-reversal of trajectories,
$d\mathcal{P}_{x_0}(\omega)$	path-probability measure for paths given x_0,
μ_0	probability distribution of initial configuration,
$d\mathcal{P}_{\mu_0}(\omega)$	path-probability measure with initial state sampled from μ_0,
$d\mathcal{P}^R$	path-probability measure with reversed protocol,
$\langle f(\omega) \rangle_{x_0} = \int d\mathcal{P}_{x_0}(\omega) f(\omega)$	expectation value of a function f,
$\langle f(\omega) \rangle_{\mu_0} = \int d\mathcal{P}_{\mu_0}(\omega) f(\omega)$	expectation value of a function f,
$\frac{d\mathcal{P}^*_{x_0}}{d\mathcal{P}_{x_0}}(\omega)$	Radon-Nikodym derivative,
$A(\omega) = -\log \frac{d\mathcal{P}^*_{x_0}}{d\mathcal{P}_{x_0}}(\omega)$	the action,
$\mu_t(x)$	time-evolved probability distribution,
$\rho(x)$	stationary distribution,
$j_\mu(x)$	probability current when in μ.

Entropy and traffic

$S(\omega) = \log \frac{d\mathcal{P}_{x_0}}{d\mathcal{P}^R_{x_T}\theta}(\omega)$	entropy flux into environment,
$S_\mu(\omega) = \log \frac{d\mathcal{P}_{\mu_0}}{d\mathcal{P}^R_{\mu_T}\theta}(\omega)$	measure of irreversibility,
$s(\mu)$	Shannon/Gibbs entropy of μ,
$\sigma(\mu)$	expected entropy production rate when in μ,
$S_{ex}(\omega) = A^R(\theta\omega) - A(\omega)$	excess entropy flux,
$\mathcal{T}_{ex}(\omega) = A^R(\theta\omega) + A(\omega)$	excess traffic,
$\tau(\mu)$	expected traffic rate when in μ.

Markov jump processes

$x, y, \ldots$	configurations,
$k_t(x, y)$	transition rates,
$\lambda_t(x) = \sum_y k_t(x, y)$	escape rate.

Diffusions

x	position,
$x^i = x_i$	components of the position,
v	velocity,
m	mass,
γ	friction coefficient,
χ	mobility,
D	diffusion coefficient,
B_t	Wiener process.

Fluctuation-dissipation

h_t	time-dependent amplitude of the perturbation,
V	perturbing potential,
Q	observable, not to be confused with heat,
$R_{QV}(t, s)$	response function,
$\tau(\omega, s)$	functional derivative of excess traffic w.r.t. h_s.

Dynamical fluctuations

p_ω	empirical occupation vector/density,
J_ω	empirical current,
μ	fluctuation of the occupations,
j	fluctuation of the current,
I	rate function or fluctuation functional,
K_f	a quantity K computed in a dynamics determined by a force f,
μ_1	small deviation of μ from ρ,
j_1	small deviation of j from j_ρ.

List of figures

Contents

Chapter 1

Introduction

In this chapter we quickly review some important concepts in thermodynamics and statistical mechanics that are relevant for the rest of this thesis. For a more thorough introduction we refer to standard textbooks. After this we introduce the reader to the realm of nonequilibrium phenomena and motivate the research discussed in this thesis.

1.1 Thermodynamics

Thermodynamics studies energy conversions between mechanical work and heat. Although several results in thermodynamics date back to the seventeenth century, the theory as we now know it had its major breakthrough in the nineteenth century, starting with groundbreaking theoretical considerations by Sadi Carnot. It is no coincidence that this happened in the century of the industrial revolution. Thermodynamics provided a deep and indispensable understanding of the principles by which engines (and refrigerators) operate, and the fundamental limits they must obey. It is also a logical starting point for our discussion. In the following pages we quickly review some aspects of thermodynamics that are relevant to the rest of this thesis; it is certainly not our goal to review the whole theory of thermodynamics.

Figure 1.1: Sadi Carnot, 1796 - 1832.

Classical thermodynamics as developed in the nineteenth century is a theory of macroscopic systems. As we now know, a macroscopic system consists of a large (10^{23}) number of particles (atoms/molecules). Thermodynamics thus describes systems with variables such as temperature, pressure, volume, etc, ignoring the microscopic details of the system (on the level of the molecules). When these variables change in time we speak of a thermodynamic process.

Equilibrium As this text is situated in nonequilibrium thermal physics, it is useful to describe what it means to be *in* equilibrium. With equilibrium we mean the following:

- Two systems are said to be in mechanical equilibrium with each other when the pressures (P) they exert on each other are equal. If the pressures were different, one system would do work on the other, causing a change in the volumes (V).
- Two systems that can exchange particles are in diffusive equilibrium when their chemical potentials (μ) are the same. If they are not then there will be a net current of particles from one system to another, causing the particle numbers (N) of the system to change.
- Two systems are in thermal equilibrium when, after being brought into thermal contact with each other, they do not exchange heat. In this case their temperatures (T) are the same. If the temperatures are not the same, there will be heat flow from one system to the other, causing a change in a new quantity, named entropy (S). (We will come back to this later).
- Two systems are in thermodynamic equilibrium if they are in mechanical, diffusive and thermal equilibrium.

Usually we do not speak about two systems but about one system and its environment, and say that a system is in equilibrium if it is in equilibrium with its environment.

These definitions already tell us that we distinguish three ways of exchanging energy between systems: work, particle exchange and heat exchange. Furthermore six variables are introduced, grouped in three pairs: P, V and μ, N and T, S. These variables are not independent. Simple systems, such as gases (or more generally pure fluids) can be described by taking one variable of each group, depending on the interaction of the system with its environment. This collection of variables is then called the (equilibrium) state of the system. For example, if the system is mechanically isolated from the environment, then its volume is fixed and can be used to describe the system. One can also fix the pressure and let the volume vary, thus allowing exchange of work, and so on. For example a system that can

exchange work and heat but not particles with its environment is best described by the variables P, N, T.

Note that V, N, S are extensive variables, i.e. they scale with the 'size' of the system (e.g. if we take two copies of the same system, it has twice the volume, twice the number of particles and twice the entropy). In contrast P, μ, T are intensive variables, i.e. they are independent of the size of the system.

There is a special class of thermodynamic processes which have the following property: if the process is run backwards, eventually the system and its environment return to the same equilibrium state they had before the original process. We call this a reversible process. Actually, for this to be true in real systems, the process should go infinitely slowly, and can in that case be described by a sequence of equilibrium states. As it turns out, many processes in real life are slow enough for this description to be a reasonable approximation.

Of course, not all quantities mentioned here have a clear intuitive meaning. Especially the quantities heat and entropy that have to do with thermal equilibrium are vague. The laws of thermodynamics give a further specification of these concepts. We discuss them here shortly, as the understanding of entropy and its role in nonequilibrium systems is very important throughout this thesis.

The first law The internal energy U of a system corresponds microscopically to the sum of all kinetic energies of the particles and all their interaction potentials. The first law of thermodynamics then dictates that a small change of the energy dU of a system can only be caused by work, heat or a particle flow:

$$dU = đQ + đW + \mu dN$$

where $đQ$ is an infinitesimal amount of heat flow into the system and $đW$ is an infinitesimal amount of work done on the system. The notation with đ stems from the fact that heat and work can't generally be expressed as the difference of a state function (a function only depending on the state of the system, not on its history). If the system contains different species of particles with chemical potentials μ_i then the last term should be replaced by $\sum_i \mu_i dN_i$.

For a system with a fixed number of particles and a fixed volume, the first law tells us that the energy of a system can only change through heat flow. In a sense this is a definition of heat: it is an energy transfer that is not work and not a particle current.

To have a better understanding of the concept of heat let us make a little thought experiment. Think of a box containing a gas (system). We imagine the box standing in a room filled with air (environment). The box is closed and has a fixed volume. Going down to the microscopic level (leaving thermodynamics for

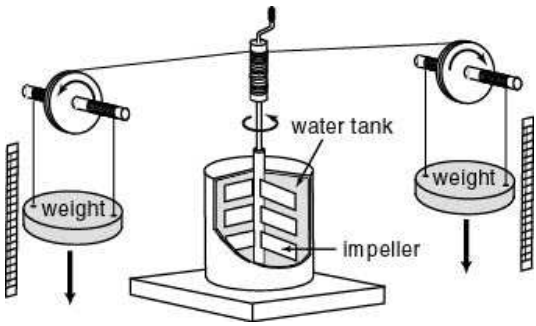


Figure 1.2: The experiment with which J.P. Joule showed that heat and mechanical work are both forms of energy transfer.

a moment), we can imagine that particles outside the box collide with the box, thus exchanging work with the box. The particles in the box also exchange work with the box. In this way energy can be transferred between the system and the environment. However when we zoom out to the macroscopic level, we can no longer see the individual collisions of all the particles, and if we do not see forces, we can't compute their work. We only see that energy is exchanged. In this way heat is defined as energy exchange due to work of forces we can't see from our macroscopic point of view. More generally heat can also be particle exchange of particles we can't see. Note that this definition is arbitrary in that it depends on the level of description of the system.

The second law This law is used to define the quantity entropy. It states that there exists a state function, called the entropy S, such that

$$\Delta S \geq \int \frac{đQ}{T}$$

where $\Delta S = S_f - S_i$ is the change in entropy of the system from the initial value S_i to the final value S_f, and the integral is over a thermodynamic process. Moreover, only for a reversible process the inequality becomes an equality, thereby exactly defining entropy changes. Apart from defining the entropy, this second law is also a restriction on the heat flow during a process. For example this law predicts that heat never spontaneously flows from a cold to a hot object. It also places a bound on engines that extract work by utilizing two reservoirs at different temperatures.

For an isolated system (i.e. $đQ = 0$), what does it mean to be 'in equilibrium'? From the second law we see that any process the system undergoes will increase its entropy (or leave it unchanged). Only when the system reaches its equilibrium

state, the entropy does not change. One can thus characterize equilibrium for an isolated system as the state with maximum entropy.

Thermodynamic potentials Thermodynamic potentials are (scalar) functions that describe the thermodynamic state of a system. From them, many relations between thermodynamic quantities can be derived. Therefore they play a role in thermodynamics comparable to the Lagrangian or Hamiltonian in mechanics. Part of this thesis discusses the use of thermodynamical potentials in nonequilibrium systems. Because of this, we provide here a quick introduction to these potentials in thermodynamics.

The first and most intuitive thermodynamic potential is the internal energy of the system: U. We can use the first law for a reversible process to write

$$dU = TdS - PdV + \mu dN$$

where PdV is the work done by the system on its environment. From this we can see that the natural variables of U are S, V and N, while the variables T, P, μ all depend on them through

$$\frac{\partial U}{\partial S} = T, \quad \frac{\partial U}{\partial V} = -P, \quad \frac{\partial U}{\partial N} = \mu$$

The internal energy is thus fixed in a system for which S, V, N are fixed and is thus the most natural potential here. Equivalently we can write the entropy $S = S(U, V, N)$ which is most natural in a system with U, V, N fixed, i.e. a totally isolated system. As we already argued, the entropy characterizes the equilibrium state of such a system because it is maximal then. Such a characterization of equilibrium is a key feature of thermodynamic potentials.

For a system with a fixed volume and particle number, but which can exchange heat with an environment which is in equilibrium at a fixed temperature, the natural variables are T, V, N. What characterizes equilibrium for such a system? To see this, we assume that the total of system plus environment is isolated. The energy of an isolated system is constant. So $dU + dU_e = 0$, where dU is the change of energy of the system, and dU_e of the environment. From the first law of thermodynamics, we see that for the environment (which is at equilibrium): $dS_e = \frac{1}{T} dU_e = -\frac{1}{T} dU$. The system is in equilibrium with its environment if the total of the two is in equilibrium. In that case the total entropy is maximal:

$$d(S + S_e) = 0, \quad d^2(S + S_e) < 0$$

Written in terms of the system, this gives:

$$d(S - \frac{U}{T}) = 0, \quad d^2(S - \frac{U}{T}) < 0$$

This defines a thermodynamic potential $S - U/T$, but a slightly different potential is more commonly used: the Helmholtz free energy $F = U - TS$. We see that it is minimal in equilibrium. This free energy for an equilibrium system is thus defined by

$$F(T, V, N) = \inf_S [U(S, V, N) - TS]$$

which means that it is a Legendre transform of the internal energy U.

Similarly the enthalpy

$$H = U + PV$$

is defined for a system that is thermally isolated but can change its volume with an environment at constant pressure. In equilibrium we have

$$H(S, P, N) = \inf_V [U(S, V, N) + PV]$$

The Gibbs free energy

$$G(T, P, N) = U + PV - TS$$

is defined for a system in contact with an environment at a constant temperature and pressure. More potentials can be defined when particle exchange with the environment is allowed.

The collection of thermodynamic potentials, all connected through Legendre transforms, form a very powerful and useful theoretical formalism for thermodynamics of equilibrium systems.

1.2 Equilibrium statistical mechanics

Statistical mechanics, also called statistical thermodynamics, is the theoretical framework that explains thermodynamics as a set of macroscopic properties of materials, starting from the microscopic properties of the individual atoms or

molecules. But it does more: it gives a more accurate description because it also describes the fluctuations from the expected macroscopic behaviour. It uses statistics/probability theory to be able to study systems that consist of a large number (10^{23}) particles. Historically, it was founded in the second half of the nineteenth century, mainly by Ludwig Boltzmann and James Clerk Maxwell, but also by Gibbs, Einstein and Planck. Their statistical mechanics is a theory of equilibrium systems. We give a quick overview of some relevant aspects:

Micro and macro In statistical mechanics one makes the difference between microstates and macrostates. A microstate (or microscopic configuration) of a system contains all microscopic information about the system, such as the positions and velocities of all its particles. A macrostate (or macroscopic configuration) gives a description of the system in terms of a few macroscopic properties, such as temperature and volume. Many microstates may correspond to the same macrostate.

Entropy One of the basic postulates of statistical mechanics is that for an isolated system, all its possible microstates are equivalent, i.e one is not more important or more probable than another. As a consequence, for an isolated system, macrostates that correspond to more microstates are more probable than macrostates that correspond to fewer microstates. This is closely connected to Boltzmann's famous definition of entropy as a statistical quantity: given a set of macrostates M, the entropy S is proportional to the logarithm of W, where W is the number of microstates corresponding to M:

$$S = k_B \log W \tag{1.1}$$

where k_B is Boltzmann's constant. For continuous microscopic configuration spaces, W is the 'volume' of the set of microstates corresponding to the macrostate. Saying that a system in equilibrium is characterized by a maximal entropy is thus the same as saying that it is characterized by the macrostates that correspond to the most microstates, i.e. the most probable macrostates. One of the major achievements of statistical mechanics is that this entropy coincides with its

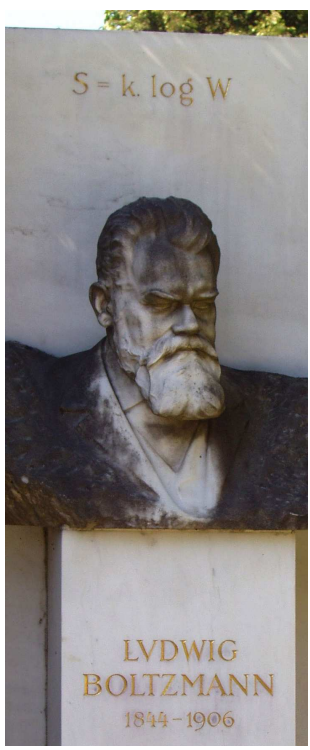


Figure 1.3: Boltzmann's gravestone.

thermodynamic counterpart mentioned above. As a testament to the significance of this formula (1.1), it was engraved on the tombstone of Boltzmann.

Ensembles The macrostate M that a system is in at any moment depends on the microstate x of the system: $M = M(x)$. Statistical mechanics explains the measured value of M from the viewpoint of the microstate x the system can be in. It does this by working with ensembles. An ensemble is a set of (a large number of) imagined copies of a system: one for each microstate the system could be in. One then has to choose a probability distribution $\rho(x)$ on this set of possible microstates. The probability of observing a certain macrostate M is then the sum (or integral) of $\rho(x)$ over all microstates that correspond to M. The statistical average of M can be computed by

$$\langle M \rangle = \sum_x \rho(x) M(x)$$

where the sum can be an integral, depending on the system. Within the theory of equilibrium ensembles the definition of entropy was provided by Gibbs:

$$S = -k_B \sum_x \rho(x) \log \rho(x)$$

which is equivalent to the Boltzmann entropy for the microcanonical ensemble (see below).

As macroscopic systems typically have 10^{23} particles, and thus a large number of microstates, the law of large numbers from probability theory dictates that the measured value of M is practically always equal to the computed average $\langle M \rangle$. Therefore, by choosing the right probability distribution ρ, one can recover thermodynamics by considering the statistical averages of macrostates. The most commonly used probability distributions are:

- The microcanonical ensemble is used for isolated systems, i.e. systems that have a fixed energy E. The set of possible microstates is thus restricted to microstates that have an energy E. Apart from that an equal probability is assigned to all those microstates.
- The canonical ensemble is used for systems that can exchange energy only in the form of heat with an environment at a fixed temperature T. The canonical distribution is

$$\rho(x) = \frac{1}{Z} e^{-\beta E(x)}$$

where Z is a normalization factor, called the canonical partition function, $\beta = \frac{1}{k_B T}$ and $E(x)$ is the energy of the system in microstate x.

- The grand canonical ensemble is used for systems that can exchange heat and particles with a reservoir at a fixed temperature T and chemical potential μ:

$$\rho(x) = \frac{1}{\mathcal{Z}} e^{\beta\mu N(x) - \beta E(x)}$$

where $\mathcal{Z}$ is a normalization factor, called the grand canonical partition function, and $N(x)$ is the number of particles of the system in microstate x.

As an example of how the link to thermodynamics is made: we explicitly substitute the distribution for the canonical ensemble in the Gibbs entropy:

$$S = k_B \log Z + \frac{\langle E \rangle}{T}$$

The average energy $\langle E \rangle$ coincides by the law of large numbers to the measured energy U, so

$$-\frac{1}{\beta} \log Z = U - TS = F$$

where F is the Helmholtz free energy of the system.

1.3 Out of equilibrium

In the end, the purpose of statistical mechanics is twofold: on the one hand to describe the macroscopic world, and derive its physical behaviour, starting from the microscopic level. On the other hand, statistical mechanics makes more predictions than thermodynamics, because it also describes the deviations (fluctuations) from the average behaviour.

Equilibrium statistical mechanics is in this sense very powerful, because it is seen that measurable quantities, using the law of large numbers, can be computed through averaging over equilibrium distributions of microstates. These are themselves expressed solely in terms of external constraints (temperature, chemical potential), and conserved quantities like energy and particle number. Going beyond thermodynamics, for example the fluctuation-dissipation theorem is a useful and well-known physical relation. (Part of this text discusses the generalization of this to nonequilibrium systems).

However, up to the moment that this text is written, there is no general paradigmatic theoretical framework that describes systems *out of* equilibrium, neither in thermodynamics nor in statistical mechanics, not even for stationary systems. This poses a problem, because many processes in real life can't be described as reversible processes, and many systems are not in equilibrium.

How to be out of equilibrium In general we distinguish two ways for systems to be out of equilibrium. First, systems that are in the process of relaxing to equilibrium: when a parameter determined by the environment changes, like temperature or volume, the system always needs some time to relax to that new equilibrium state. For example, when a hot cup of coffee is placed in a room at room temperature, heat will start to flow from the coffee to the air of the room (we ignore heat exchange by radiation with the walls of the room). After some time the coffee and the air will relax to a new equilibrium state: the coffee cools down to room temperature. Before that time, however, the coffee is not in equilibrium with the room, and the process is not reversible. If it was reversible, then we would not be surprised to observe heat to flow spontaneously from the air to the coffee, heating it up again. Sometimes the external parameters change continuously and fast enough such that the system never has enough time to relax to equilibrium. Think for example of combustion engines.

Secondly there are systems that are driven from equilibrium by what we call thermodynamic forces. Here we distinguish three subgroups, corresponding to thermal, diffusive and mechanical exchanges of energy:

- Systems in contact with parts of the environment at different temperatures. For example a wall of a house: on one side it is in contact with the warm air inside, and on the other side it is in contact with cold air. In such systems there are constantly heat currents. A part of the environment that is in thermal contact with the system, is called a heat bath. The thermodynamic force here is the temperature difference.
- Systems in contact with parts of the environment at different chemical potentials. Think of a cell membrane with a bigger particle density on the inside than on the outside. In such systems one sees particle currents. A part of the environment that is in diffusive contact with the system is called a particle reservoir. The thermodynamic force here is the chemical potential difference.
- Systems under the influence of mechanical nonconservative forces. A nonconservative force is a force which is not the derivative from a potential. Think of the pressure difference that makes water run out of a tap.

To simplify this second group of systems, one usually assumes that the parts of the environment in contact with the system are and stay in equilibrium during the process. For this one should assume that there is no interaction between the different parts of the environment, and that the interaction between system and environment is weak, i.e. changes in each part of the environment are small enough such that it relaxes fast enough to equilibrium for our description of the system. One also assumes that the heat baths and particle reservoirs are big, such that during the process their temperatures and densities do not change measurably. Systems under such conditions can often relax to a stationary state. This is a state in which the variables (macrostates) with which we describe the system do not change in time. Such variables include heat or particle currents in contrast to equilibrium systems. One of the goals of nonequilibrium thermodynamics is to describe such stationary regimes and correctly predict the directions and sizes of their currents.

Dynamics An important property of nonequilibrium systems, in contrast to equilibrium, is that time plays an important role. First of all because nonequilibrium processes are irreversible: an arrow of time is introduced. The question of how irreversibility emerges when going from the reversible microscopic world to the macroscopic world was already discussed by Boltzmann himself.

Apart from that, it seems clear that nonequilibrium systems are by their very nature dynamical. This is because systems are either out of equilibrium because they are driven from equilibrium and are in some stationary regime in which there are particle or heat currents present, or because they are in the process of relaxing to equilibrium (or to a stationary state). This means that not only the microstates themselves are important, but also the way in which they change in time. The dynamics of the process should enter our (theoretical) description.

Statistical mechanics of stochastic processes: the mesoscopic level Ideally, one would therefore like to answer physical questions about average behaviour and fluctuations, by starting from the microscopic Hamiltonian dynamics. However, this is often just impossible. This is not necessarily a disaster, because we expect (hope) that not every detailed aspect of that dynamics is relevant for the macroscopic world. Therefore most models used in nonequilibrium statistical mechanics are already reduced descriptions, meaning that they do not contain all information of the microscopic world. For example, one usually reduces the description of the environment of the system to some parameters as temperature or chemical potential. On the other hand, to be able to describe fluctuations from the typical macroscopic behaviour, a more detailed description than the macroscopic one is needed. We then say that we are working on the mesoscopic level.

One should thus think of a mesoscopic model as describing a small system, for which the description is not detailed enough to be Hamiltonian, but is not big enough for the law of large numbers to apply. Fluctuations around the statistical averages are important. As a consequence a model describing a mesoscopic system can be a stochastic process. A stochastic process is therefore a very important tool in nonequilibrium statistical processes. An important part of the present day research is therefore committed to finding 'recipes' for defining stochastic processes that are physically relevant. One way to do this is via the local detailed balance assumption, which will be discussed in the next chapters, and is used throughout this text.

Applications As said before, many processes in everyday life are not reversible, and many quantities of interest are some form of current, which is not found in equilibrium systems. On the macroscopic scale we find important examples in ecology and meteorology. In meteorology, we see that the temperature at the poles is on average much lower than at the equator. That is responsible for major energy currents in the earth's weather system. On a smaller scale one is interested in predicting local weather, like temperature changes and air flows (wind), given observational data concerning high and low pressure areas etc. In ecology one is interested in the energetic and material (food) flows between different species of organisms.

On a much smaller scale we see that mesoscopic models are not only useful as a step towards macroscopic theories, but also as a description of interesting but very small systems. In recent years the nonequilibrium world of small systems has become more accessible experimentally, opening and expanding research areas as biophysics and nanotechnology. In biophysics a lot of attention nowadays is going to e.g. the dynamics of DNA and RNA, transport of ions through cell membranes, molecular motors, etc. In nanotechnology the electrical current in extremely small devices or parts of devices is central. Finally, in everyday life there are many finite networks, like transport and communication networks, on which processes occur which are well described on the mesoscopic scale.

A lot of interesting and useful research has been done and is being done in these areas, where explicit models (stochastic processes) are examined to derive results for each specific area of interest. However, this is not the goal of the research reported in this thesis. Instead, we try to find general physical structures of nonequilibrium statistical mechanics. Such a scheme should in the long run and together with many other contributions lead to a better understanding of nonequilibrium physics. Such an understanding should then give useful predictions applicable to the specific areas of interest mentioned above.

1.4 Outline and preview of the results

Part I: Stochastic processes Because of the lack of a general theory, it may come as no surprise that the present day research into nonequilibrium systems focuses on simple physical systems. It is then important to define a stochastic process that models it correctly. The first part of this thesis therefore contains an introduction to the stochastic models that are used in the rest of the text, focusing on the relevant aspects for understanding the next two parts.

Throughout this text, we restrict ourselves to classical systems, meaning that we do not consider relativistic or quantum effects. (Quantum mechanics only enters in some specific models in the discretization of configuration space).

Part II: Fluctuation-dissipation relations In this part of the thesis, we investigate how a system responds to a perturbation, namely a small change in its energy. The central object that summarizes this is the response function. To be more precise, we denote the energy of the system in configuration x by $U(x)$. As a perturbation this energy is changed by the addition of a potential: $U \to U - h_t V$, where the time-dependent function h_t is the amplitude of the perturbation. We restrict our possible class of systems to those in which a small amplitude h_t only has small effects (excluding for example the regime of phase transitions). Then we can write the expectation of an observable Q in the perturbed system as the expectation in the unperturbed system plus a small correction:

$$\langle Q(x_t)\rangle^h \approx \langle Q(x_t)\rangle^0 + \int_0^t ds h_s R_{QV}(t,s)$$

This defines the response function $R_{QV}(t,s)$. In equilibrium systems it is known that this response function can be written in terms of a correlation function in the unperturbed system (a system in equilibrium at inverse temperature β):

$$R_{QV}(t,s) = \beta \frac{\partial}{\partial s} \langle Q(x_t) V(x_s)\rangle^0$$

This is called the fluctuation-dissipation theorem. Our contribution has been to discuss the form of the response function for systems out of equilibrium. We found that the response function can generally be written as the sum of two correlation functions. One correlation function is linked to heat dissipation into the environment, and thus to entropy changes. The other correlation function has to do with a quantity which we call traffic, which describes in a sense the activity of the system. The main results of this part are summarized by formulae (5.6), (5.14) and (6.7). These results are then applied to several explicit examples for which simulations have provided some visualization and verification.

The research discussed in this part was reported before in [2, 3, 4].

Part III: Dynamical fluctuations In this part of the thesis we use the theory of large deviations to examine certain fluctuations in systems out of equilibrium. For equilibrium statistical mechanics, large deviation theory provides a natural mathematical framework especially for the thermodynamic potentials discussed above. It is therefore useful to investigate that theory also for nonequilibrium systems.

In this theory we consider the empirical occupation density p (which describes the fraction of time the system spends in each configuration) and the empirical current J (describing the changes of configuration the system makes) as observables. The probability is then considered that these observables take on some given values $p \approx \mu$ and $J \approx j$. Such a probability often shows an exponential decay, whenever μ and j do not correspond to the typical values of the observables:

$$\mathrm{Prob}(p \approx \mu, J \approx j) \approx e^{-TI(\mu,j)}$$

where the duration T of the process is very large, and $I(\mu, j)$ is called the rate function.

The questions in this part of the thesis are then: how to compute the rate function, and what are the physical quantities governing its form? As in the previous part, entropy and traffic are the main ingredients. Moreover, the rate function $I(\mu, j)$ is explicitly computed for the classes of models considered and expressed in terms of entropy and traffic, see (10.25) and (11.10). The rate function for the occupations can be expressed entirely in terms of traffic. We also show that this traffic can be seen as a thermodynamic potential for currents. Finally, for the close-to-equilibrium regime, known variational principles such as the minimum entropy production principle are recovered.

The research discussed in this part was published before in [70, 71].

Part I

Stochastic processes

"Observe what happens when sunbeams are admitted into a building and shed light on its shadowy places. You will see a multitude of tiny particles mingling in a multitude of ways... their dancing is an actual indication of underlying movements of matter that are hidden from our sight... It originates with the atoms which move of themselves. Then those small compound bodies that are least removed from the impetus of the atoms are set in motion by the impact of their invisible blows and in turn cannon against slightly larger bodies. So the movement mounts up from the atoms and gradually emerges to the level of our senses, so that those bodies are in motion that we see in sunbeams, moved by blows that remain invisible."

Lucretius, *De rerum natura*, ca. 60 BC.

Chapter 2

Describing stochastic processes

Stochastic processes are a useful tool to explore nonequilibrium statistical mechanics. The basic ingredient here is not the microstate but the trajectory, i.e. the sequence of states the system visits during the process. In this chapter we therefore discuss how to work with such trajectories in statistical physics. To lift the mathematical models to a more physical level, the local detailed balance assumption is made and explained. As we will see, some well-known general physical results are a direct consequence of this assumption.

2.1 Stochastic processes

Mathematically we describe a system by its configuration or state. This state contains all information of the system that is relevant for our description. Let us denote it by x, being an element of the configuration space Ω. This x can be for example the vector of positions and momenta of a gas, or the set of spin-configurations of a magnet, or maybe just the position of one particle submerged in a fluid. Consequently, the set Ω can be continuous or discrete, unbounded, compact or even finite.

In the course of time the configuration of the system will change, i.e. the configuration depends on time: $x = x_t$. How it changes, is determined by the dynamics of the process. In classical mechanics, the dynamics at the microscopic level is Hamiltonian and therefore deterministic. Indeed, when all positions and momenta of the particles of a gas in an isolated container are given at a time t, one can in principle determine the positions and momenta at any later time.

Unfortunately, this is not always possible. Most of the times the system under consideration is not isolated, but interacts with some environment of which the exact configuration is unknown. Moreover it is usually even impossible to measure the exact configuration of the system itself. Because of such factors, uncertainty enters, and must enter our mathematical description. This gives rise to stochastic processes, and the configuration x_t of the system at each time becomes a stochastic variable X_t. A stochastic process is then the sequence of these stochastic variables:

$$\{X_t | 0 \leq t \leq T\}$$

When the history of the system is given, i.e. x_t for $t \in]-\infty, T]$, the stochastic dynamics of the process gives a probability to find the system in a configuration y at a later time.

Again one has to be more modest, as the complete history of a system is rarely known. A widely used and usually well-working approximation is the Markov approximation. This consists of the assumption that the probability of finding the system at a later time in some configuration, only depends on the configuration of the present, not the past. Informally, for a sequence of times $t_1 < t_2 < \ldots < t_n < t$:

$$Prob(x_t = x | x_{t_n} = y_n, \ldots, x_{t_1} = y_1) = Prob(x_t = x | x_{t_n} = y_n) \tag{2.1}$$

A Markov process is a stochastic process for which the Markov assumption (2.1) is valid. Note that Hamiltonian dynamics, governing the evolution of the state (x, v), although deterministic, also satisfy this assumption. In the rest of this text we always assume our processes to be Markovian. In many cases the dynamics of the Markov process have the property that

$$Prob(x_t = x | x_s = y) = Prob(x_{t-s} = x | x_0 = y) \qquad \forall s < t$$

In such cases we speak of time-independent dynamics (or time-homogeneous dynamics). However, this is not always the case in this thesis.

For the reader who is unfamiliar with probability theory and stochastic processes, we refer to [46] for a thorough introduction.

A particle in a box As an easy introductory example, consider a box with volume V filled with air. In this box we place one charged test-particle. We also apply an electric field to move the particle in our favourite directions. If we knew all the positions and velocities of all the particles in the box, in principle the trajectory of test-particle could be computed with Hamilton's equations. Practically, this

is not possible, but suppose that in some way the position of the test-particle can be measured at time intervals of 1 second, the first measurement being at time zero. The precision of the measurement is limited: one can only determine the position up to a small volume ΔV. Therefore we divide the box into $N = V/\Delta V$ parts, labelled by $x = 1, 2, \ldots$. Given the configuration (position) x after n seconds, denoted by x_n, the position x_{n+1} is not determined exactly. There are usually several possible positions that the particle can be in, each having a probability determined by the temperature of the air, the electric and magnetic fields, etc. These probabilities can also depend on the previous positions of the particle. However, let us assume that the interaction with the air molecules is sufficiently strong compared to the influence of the electromagnetic fields and inertial effects, such that the particle loses its memory quickly enough for the Markov assumption to be valid.

Having this physical example in our mind, let us write the probabilities of the successive positions (transition probabilities) as follows:

$$P_n(x, y) = Prob(x_{n+1} = y | x_n = x)$$

Mathematically, this defines what is called a Markov chain, which is the simplest example of a Markov process. In the next two chapters we discuss two other classes of Markov processes: Markov jump processes and diffusions, but in this chapter we use these simple Markov chains to illustrate the introduced concepts.

The $N \times N$ matrix P_n is called the transition matrix at time n, and it determines the dynamics of the process. Note that we should have that $\sum_y P_n(x, y) = 1$ for every n, where the sum is of course over all configurations. When the electromagnetic fields are constant in time, the dynamics are time-homogeneous: $P_n = P$. When the fields are time-dependent, so are the dynamics. The time-dependence of the fields is determined by some parameter λ_t which continuously changes in the time t, so $P_n = P_{\lambda_n}$. This λ_t is controlled externally, and we imagine it to be deterministic. It is called the protocol.

2.2 Trajectories

A realization of a stochastic process is called a trajectory or path. We denote it by $\omega = (x_t)_{0 \leq t \leq T}$. If the dynamics of the stochastic process are given, one can in principle compute the probability measure on such paths ω. Suppose that the configuration of the system at time zero x_0 is given. We denote by $d\mathcal{P}_{x_0}(\omega)$ the path-probability measure of ω given x_0. With it we can compute the expectation

values of observables: take any observable $f(\omega)$ that depends on the trajectory. Then its expectation value is

$$\langle f(\omega)\rangle_{x_0} = \int d\mathcal{P}_{x_0}(\omega) f(\omega) \tag{2.2}$$

where one integrates (or sums, depending on the model) over all possible trajectories that start from x_0 at time zero. More generally, given a probability distribution (or density) μ_0 of initial configurations, we denote the probability measure of a path by $d\mathcal{P}_{\mu_0}(\omega) = \mu_0(x_0) d\mathcal{P}_{x_0}(\omega)$ and the expectation value of $f(\omega)$ by

$$\langle f(\omega)\rangle_{\mu_0} = \int d\mathcal{P}_{\mu_0}(\omega) f(\omega)$$

where integration/summation is now over all possible trajectories.

Basic probabilistic rules dictate that the path-probability measure should be normalized:

$$\int d\mathcal{P}_{\mu_0}(\omega) = 1 \tag{2.3}$$

Moreover, if we split the trajectories into $\omega_1 = (x_t)_{0\leq t\leq s}$ and $\omega_2 = (x_t)_{s<t\leq T}$, for any s in the interval $[0, T]$, then the path-probability measure of paths ω_1, which we just denote by $d\mathcal{P}_{\mu_0}(\omega_1)$, is found by integrating over all ω_2:

$$d\mathcal{P}_{\mu_0}(\omega_1) = \int_{\omega_2} d\mathcal{P}_{\mu_0}(\omega) \tag{2.4}$$

This is an important property. For example, in many computations in this thesis we take expectation values of state functions, evaluated at some time s. This gives

$$\langle f(x_s)\rangle_{\mu_0} = \int d\mathcal{P}_{\mu_0}(\omega) f(x_s) = \int d\mathcal{P}_{\mu_0}(\omega_1) f(x_s)$$

i.e. we only need to take an average over all paths in the interval $[0, s]$.

Particle in a box For Markov chains, a trajectory is determined by the successive configurations: $\omega = (x_0, x_1, \ldots, x_m)$, where m corresponds to the time T. The

probability of a trajectory ω is just $\mathcal{P}_{x_0}(\omega) = \prod_{n=0}^{m-1} P_n(x_n, x_{n+1})$. If we take for example a function $f(x)$ evaluated at time m, then

$$\langle f(x_m)\rangle_{x_0} = \sum_{x_1,\ldots,x_m} \left[\prod_{n=0}^{m-1} P_n(x_n, x_{n+1})\right] f(x_m)$$

2.3 Probability distribution of states

The probability of finding the system in a configuration x at time t is denoted by $\mu_t(x)$, the time-evolved distribution (actually in the case of a continuous configuration space, $\mu_t(x)$ is a probability density). It can easily be written in terms of expectation values:

$$\mu_t(x) = \langle \delta(x_t - x)\rangle_{\mu_0} \tag{2.5}$$

using $\delta_{x_t,x}$ instead of $\delta(x_t - x)$ for discrete state spaces. Equivalently, for any state function $f(x)$:

$$\langle f(x_t)\rangle_{\mu_0} = \int dx \mu_t(x) f(x) \tag{2.6}$$

The integral becomes a sum for discrete state spaces. In systems with time-independent dynamics, there often exists a distribution which does not change under the time-evolution. This is called the stationary distribution. If the system is in the stationary distribution, it is said to be in the 'steady state'. We always denote the stationary distribution with ρ:

$$\rho(x) = \langle \delta(x_t - x)\rangle_{\rho} \qquad \forall t$$

Under certain conditions on the dynamics, it can be proven that such a distribution exists, is unique, and that all distributions μ_t converge to it in the long time limit (the system relaxes to the stationary distribution). Unless stated otherwise, we always assume this to be true for time-independent dynamics. See [46] for a rigorous treatment.

The existence of such a stationary measure implies ergodicity, meaning that for trajectories ω and an arbitrary state function g:

$$\frac{1}{T}\int_0^T dt g(x_t) \to \int dx \rho(x) g(x) \qquad \text{for } T \to \infty \tag{2.7}$$

almost surely. More precisely: the probability that the system follows a trajectory that satisfies (2.7) is equal to one.

Particle in a box For a Markov chain we get:

$$\mu_m(x) = \sum_{x_0,\ldots,x_{m-1}} \mu_0(x_0) \left[\prod_{n=0}^{m-2} P_n(x_n, x_{n+1})\right] P_{m-1}(x_{m-1}, x)$$

Note that the distribution one time-step later can be written as $\mu_{m+1}(x) = \sum_y \mu_m(y) P_m(y, x)$, and thus:

$$\mu_{m+1}(x) - \mu_m(x) = \sum_y [\mu_m(y) P_m(y, x) - \mu_m(x) P_m(x, y)]$$

In the case that $P_n = P$, if there is a stationary distribution it has to satisfy

$$0 = \sum_y [\rho(y) P(y, x) - \rho(x) P(x, y)]$$

2.4 The action

Suppose now that one has two stochastic processes, each having the same configuration space but a different dynamics. The action gives a way of switching between expectation values computed in those dynamics. This is essential to the rest of this text.

We denote the path-probability measures of the processes by $d\mathcal{P}$ and $d\mathcal{P}^*$ respectively. Suppose that $d\mathcal{P}^*$ is absolutely continuous with respect to $d\mathcal{P}$. This means that $\int_M d\mathcal{P}^*(\omega) = 0$ for any set of trajectories M for which $\int_M d\mathcal{P}(\omega) = 0$. Then one has

$$\langle f(\omega)\rangle^*_{\mu_0} = \int d\mathcal{P}_{\mu_0}(\omega) \frac{d\mathcal{P}^*_{\mu_0}}{d\mathcal{P}_{\mu_0}}(\omega) f(\omega) = \left\langle f(\omega) \frac{d\mathcal{P}^*_{\mu_0}}{d\mathcal{P}_{\mu_0}}(\omega) \right\rangle_{\mu_0} \tag{2.8}$$

The quantity $\frac{d\mathcal{P}^*_{\mu_0}}{d\mathcal{P}_{\mu_0}}(\omega)$ is called the Radon-Nikodym derivative between the two processes [83]. In words: it is the probability of a path in one dynamics divided by the probability of the same path in another dynamics. It is common to write this in the following way:

$$\frac{d\mathcal{P}^*_{\mu_0}}{d\mathcal{P}_{\mu_0}}(\omega) = e^{-A(\omega)} \tag{2.9}$$

where $A(\omega)$ is called the 'action.' For Markov processes it is independent of μ_0.

Normalization of the path-probability measure (2.3) tells us that

$$\left\langle e^{-A(\omega)} \right\rangle_{\mu_0} = \langle 1 \rangle^*_{\mu_0} = 1 \tag{2.10}$$

for any initial distribution μ_0.

Particle in a box For a Markov chain the action is easily computed to be

$$A(\omega) = \sum_{n=0}^{m-1} \log \frac{P_n(x_n, x_{n+1})}{P^*_n(x_n, x_{n+1})}$$

2.5 Time-reversal

To be able to talk about time-reversibility we first define an operator θ which acts as a time-reversal on trajectories. This means that it reverses all trajectories, also taking care that velocities change sign under time-reversal:

$$\omega = (x_t)_{0 \le t \le T} \mapsto \theta\omega = (\pi x_{T-t})_{0 \le t \le T}$$

where π changes signs of velocities. Moreover, if the dynamics of the process is time-dependent, we also reverse this time-dependence. For example if the dynamics is governed by a force $F_t(x)$, then time-reversal changes this to $F_{T-t}(x)$. This means that the path-probability measure will also change. Therefore, whenever we deal with time-dependent dynamics, we denote the new path-probability measure by $d\mathcal{P}^R$, (the R stands for 'reversed'). These reversals together make the time-reversal that intuitively corresponds to playing the movie of a process backwards.

Let us consider the following measure of irreversibility, defined in terms of a Radon-Nikodym derivative:

$$S_\mu(\omega) = \log \frac{d\mathcal{P}_{\mu_0}}{d\mathcal{P}^R_{\mu_T}\theta}(\omega) \tag{2.11}$$

where μ_0 is an arbitrary probability distribution from which the initial state of the system is sampled. For the time-reversed process the initial distribution is taken to be μ_T, which is defined (given μ_0) through (2.5). This quantity $S_\mu(\omega)$

is a measure of irreversibility of a certain trajectory ω. It is the probability of ω divided by the probability of its time-reversed twin in the time-reversed dynamics. Let us rewrite (2.11) as follows:

$$S_\mu(\omega) = \log\left(\frac{\mu_0(x_0)}{\mu_T(x_T)}\frac{d\mathcal{P}_{x_0}}{d\mathcal{P}^R_{x_T}\theta}(\omega)\right) =: \log\frac{\mu_0(x_0)}{\mu_T(x_T)} + S(\omega) \tag{2.12}$$

which defines the quantity $S(\omega)$. Note that for Markov processes $S(\omega)$ does not depend on the initial distribution of the system.

Particle in a box For Markov chains, time-reversal of the path gives $\theta\omega = (x_m, \ldots, x_1, x_0)$, as long as the configurations do not contain velocities, like in the example of the particle in a box. The reversal of the dynamics gives transition probabilities $P_n \to P_{m-n-1}$. Together they give

$$S(\omega) = \sum_{n=0}^{m-1} \log\frac{P_n(x_n, x_{n+1})}{P_n(x_{n+1}, x_n)} \tag{2.13}$$

2.6 Equilibrium

The most general way of defining equilibrium is that systems in equilibrium are stationary and time-reversible. The first condition means that the probability of finding the system in a state x does not depend on time, i.e. it is given by the stationary distribution $\rho(x)$ (the dynamics have to be time-independent of course). The second condition means that when we are shown a movie of an equilibrium process, we can not tell if the movie is played normally or backwards. More mathematically: $S_\rho(\omega)$ as defined in (2.11) with $\mu_0 = \mu_T = \rho$ is zero for any trajectory ω:

$$S_\rho(\omega) = \log\frac{d\mathcal{P}_\rho}{d\mathcal{P}^R_\rho\theta}(\omega) = 0 \tag{2.14}$$

This has as an immediate consequence (see (2.12)) that

$$S(\omega) = \log\rho(x_T) - \log\rho(x_0) \tag{2.15}$$

We can use our knowledge of equilibrium systems (see Section 1.2) to see what this means physically. E.g. for a system in contact with a heat bath in equilibrium at inverse temperature β, ρ is given by the canonical distribution. We get:

$$S(\omega) = -\beta[U(x_T) - U(x_0)] \tag{2.16}$$

with $U(x)$ the energy of the system in configuration x. The change of energy of the system is equal to minus the change of energy of the environment. It becomes more clear what this quantity S is when we consider a system in contact with a particle reservoir at inverse temperature β and chemical potential μ. For this ρ is the grand canonical distribution, and we get

$$S(\omega) = -\beta[U(x_T) - U(x_0)] + \beta\mu[N(x_T) - N(x_0)] \tag{2.17}$$

This is exactly the change of entropy in the environment, stated in terms of the configurations of the system (up to a factor of k_B, Boltzmann's constant, which we set to one for notational simplicity). This means that $S(\omega)$ can be interpreted as the entropy flux from the system to its environment during the trajectory ω. In the following section we generalize the connection between entropy flux and $S(\omega)$ to nonequilibrium processes.

Here we would like to point out that there is a fundamental difference between a system in equilibrium and a system with an equilibrium dynamics. As we can see from (2.12), $S(\omega)$ is independent of the initial distribution μ_0, it only depends on the dynamics of the process. So if the system has an equilibrium dynamics, (2.15) holds for any ω, even if the system is not in the stationary (equilibrium) distribution. This is also called a detailed balanced dynamics. Only if the system is also stationary, then (2.11) is zero for any ω and we say that the system is in equilibrium.

Particle in a box For a Markov chain to be in equilibrium, (2.14) must hold for any path. It should therefore hold at least for paths that last only one step: $\omega = (x_0, x_1)$. So a necessary condition for equilibrium is:

$$0 = \log\left(\frac{\rho(x_0)P(x_0, x_1)}{\rho(x_1)P(x_1, x_0)}\right)$$

for any x_0, x_1, which is equivalent to

$$\rho(x_0)P(x_0, x_1) = \rho(x_1)P(x_1, x_0)$$

This is called the detailed balance condition. One can easily check that it is also a sufficient condition for an equilibrium dynamics.

2.7 Irreversibility and entropy: local detailed balance

For a very large class of nonequilibrium systems we will assume that

$$S(\omega) = \quad \textbf{the entropy flux from the system into the environment.}$$

This assumption is called 'local detailed balance,' because it is based on the assumption that locally in time and in space the system has a dynamics that is detailed balanced. This assumption is restricted to the case that the reservoirs only interact with the system, not with each other. Moreover the coupling between system and reservoirs should be sufficiently weak and the reservoirs sufficiently big, such that the reservoirs stay in equilibrium throughout the process. The local detailed balance assumption can be seen as a restriction on the possible mathematical models that we use to describe the system. More constructively, it gives a partial recipe to write down models that correspond to the physical world. Partial, because local detailed balance does not fully specify the dynamics. The ideas behind this assumption have been used already for some time, see e.g. [62], but it was first called local detailed balance in [56]. A more rigorous treatment of this assumption was then later done in [66, 73, 74] and mostly in [67]. Local detailed balance is central to the results discussed in this thesis, and we always assume it to be true.

Let us give a heuristic motivation why such an assumption is reasonable. As said in Section 1.3, there are several ways in which a system can be driven away from equilibrium. Let us first consider the case of a system in contact with two heat baths which are themselves in equilibrium at different temperatures β_1 and β_2. The baths do not interact with each other, only with the system. Microscopically the system never interacts with the two baths at the exact same time. Imagine therefore that we let the system run for only a very short time-interval $[0, T]$. This interval is so short that the system interacts only with heat bath 1 within $[0, t]$, an only with heat bath 2 in $[t, T]$. (A more elaborate discussion of such a process can be found in [53, 54]). By its mathematical definition (2.12), $S(\omega)$ is then just the sum of contributions from $[0, t]$ and $[t, T]$. As the dynamics are detailed balanced in each time interval separately (there is interaction with only one heat bath at a fixed temperature), we can use the results of the last paragraph to see that

$$S(\omega) = -\beta_1[U(x_t) - U(x_0)] - \beta_2[U(x_T) - U(x_t)] = \beta_1 Q_1 + \beta_2 Q_2$$

where Q_1 and Q_2 denote the heat fluxes into heat bath 1 and 2 respectively. Therefore, $S(\omega)$ is the change in entropy of the environment due to the process of the system, i.e. the entropy flux from the system into the environment. Given an arbitrary time-interval, it is therefore reasonable to assume that we can split this

interval into many small intervals which are short enough so that we can assume the dynamics to be detailed balanced in each interval separately: we assume detailed balance 'locally in time'. Adding all the contributions to $S(\omega)$ from these small intervals gives us that

$$S(\omega) = \beta_1 Q_1 + \beta_2 Q_2$$

Using the same reasoning we can expand this formula to the case in which the system interacts with several heat baths and particle reservoirs. The conclusion stays the same: the quantity $S(\omega)$ is the entropy flux from the system into the environment.

Finally, a system can be driven from equilibrium by some nonconservative forcing. Let us illustrate this last class of systems with an example:

Particle in a box Consider the following example: the box in which the charged particle resides is surrounded by an environment at equilibrium at a single temperature (air or water). The air inside the box is also at equilibrium at that temperature. Suppose the box has the shape of a thin torus, with a circumference L, which we divide into N segments of length d (d would be the error of our measurement). The configuration space is thus a ring with N sites, labelled $x = 1, 2, \ldots, N, N+1 \equiv 1$. We assume that the particle can maximally move one site to the left or right during one time step. An electric field E (constant in time) is applied on the box. As a consequence, the particle gains an energy $E(x, y) \cdot d$ when going to the left or right ($y = x \pm 1$). Note that $E(x, y) = -E(y, x)$.

First of all, suppose that the electric field is of the form $E(x, y) \cdot d = U(y) - U(x)$, meaning that $U(x)$ is the energy of the particle at site x. In this case the forcing is conservative, and the dynamics of the system should satisfy the detailed balance condition:

$$\frac{P(x, y)}{P(y, x)} = e^{-\beta[U(y)-U(x)]} = e^{\beta E(x,y)\cdot d}$$

If we restrict our attention to only two sites x, y, it is impossible so say if the electric field is conservative or not. For a nonconservative force we therefore assume that detailed balance holds locally (i.e. between each pair of states separately):

$$\frac{P(x, y)}{P(y, x)} = e^{\beta E(x,y)\cdot d}$$

The quantity $S(\omega)$ is then given by (2.13):

$$S(\omega) = \sum_{n=0}^{m-1} \log \frac{P(x_n, x_{n+1})}{P(x_{n+1}, x_n)} = \beta \sum_{n=0}^{m-1} E(x_n, x_{n+1}) \cdot d$$

which is the entropy production in the environment (the entropy flux into the environment).

In general, note that the word 'conservative' is a global property. If one considers a small enough region of the configuration space, one can always define a local potential from which the force is derived. So restricting our observation momentarily to this small region, we assume that the dynamics of the system is detailed balanced (detailed balance locally in space), and still conclude that $S(\omega)$ is the entropy flux into the environment. Returning to the global picture, the sum of all these small regions gives then the total entropy flux into the environment. The only difference with a real detailed balanced dynamics is that this entropy flux is no longer the difference of a state function, but depends on the whole path ω.

2.8 Average entropy

It turns out that we can also attach a physical meaning to $S_\mu(\omega)$, as defined in (2.11), when the average is taken. This means again an average over all possible trajectories $\omega = (x_t)_{0 \le t \le T}$.

$$\langle S_\mu \rangle_{\mu_0} = \left\langle \log \frac{\mu_0(x_0)}{\mu_T(x_T)} \right\rangle_{\mu_0} + \langle S(\omega) \rangle_{\mu_0}$$

The second term on the right-hand side is the average entropy flux into the environment. The first term can be rewritten, using the fact that for any state function f, we have that $\langle f(x_t) \rangle_{\mu_0} = \int dx \mu_t(x) f(x)$ (where the integral becomes a sum for Markov jump processes). We see that the average of S_μ becomes

$$\langle S_\mu \rangle_{\mu_0} = s(\mu_T) - s(\mu_0) + \langle S(\omega) \rangle_{\mu_0} \tag{2.18}$$

with $s(\mu)$ the Shannon entropy of the distribution μ

$$s(\mu) = -\int dx \mu(x) \log \mu(x)$$

(again up to a factor of k_B). When μ is an equilibrium distribution, the Shannon entropy is equal to the Gibbs entropy, which is in equilibrium statistical mechanics

the (physical) entropy of the system. Therefore we still call it here the entropy of the system, making the average of S_μ equal to the average of the total entropy change in the world due to the process of the system.

The average entropy change is positive. This is a direct consequence of its definition,

$$\left\langle \frac{d\mathcal{P}^R_{\mu_T}\theta}{d\mathcal{P}_{\mu_0}}(\omega) \right\rangle_{\mu_0} = \int d\mathcal{P}_{\mu_0}(\omega) \frac{d\mathcal{P}^R_{\mu_T}\theta}{d\mathcal{P}_{\mu_0}}(\omega) = 1$$

and Jensen's inequality:

$$1 = \left\langle e^{-S_\mu} \right\rangle_{\mu_0} \geq e^{-\langle S_\mu \rangle_{\mu_0}}$$

2.9 Entropy and traffic

In (2.9) we wrote the Radon-Nikodym derivative, which compares the path-probability measures of two different dynamics. This defined an action $A(\omega)$. Similarly, we can define an action for the time-reversed process:

$$\frac{d\mathcal{P}^{*,R}_{\mu_0}}{d\mathcal{P}^R_{\mu_0}}(\theta\omega) = e^{-A^R(\theta\omega)}$$

As an immediate consequence of the local detailed balance assumption, we see that

$$A^R(\theta\omega) - A(\omega) = S^*(\omega) - S(\omega) = S_{ex}(\omega) \tag{2.19}$$

In words: the time-antisymmetric part of the action A is equal to the difference in entropy fluxes in the two dynamics. We call it an excess entropy flux. Here we clearly see the limits of local detailed balance: it only specifies the time-antisymmetric part of the action. And, in contrast to equilibrium dynamics, the time-symmetric part does play an important role in statistical physics out of equilibrium. Therefore it has been proposed [11, 22, 68] to examine the following quantity more closely:

$$\mathcal{T}_{ex}(\omega) = \mathcal{T}^*(\omega) - \mathcal{T}(\omega) = A^R(\theta\omega) + A(\omega)$$

The quantity $\mathcal{T}(\omega)$ is called traffic [69, 71, 70], and the time-symmetric part of the action is therefore an excess traffic. The action A can thus be written as

$$A(\omega) = \frac{\mathcal{T}_{ex}(\omega) - S_{ex}(\omega)}{2}$$

and the Radon-Nikodym derivative:

$$\frac{d\mathcal{P}^*_{\mu_0}}{d\mathcal{P}_{\mu_0}}(\omega) = e^{\frac{S_{ex}(\omega)-\mathcal{T}_{ex}(\omega)}{2}} \tag{2.20}$$

The physical and operational meaning of traffic is up to now not completely clear. From the research discussed in this thesis, it seems to be of great importance in results of nonequilibrium statistical physics.

2.10 General physical results

Using as an assumption only local detailed balance, a number of well-known results can be very generally derived. We conclude this chapter with these derivations.

2.10.1 Fluctuation theorem

A recent and celebrated result of out-of-equilibrium statistical physics is the fluctuation theorem for entropy fluxes [38, 39, 42], valid for systems with a time-independent dynamics:

$$\lim_{T\to\infty} \frac{1}{T} \log\left(\frac{P\left(S(\omega) = T\sigma\right)}{P\left(S(\omega) = -T\sigma\right)}\right) = \sigma \tag{2.21}$$

where $P\left(S(\omega) = T\sigma\right)$ is the probability density that the entropy flux is equal to a value $T\sigma$. We use local detailed balance to relate entropy fluxes to the path-probability measure. With this the derivation is not difficult. We first consider the probability density for $S_\rho(\omega)$ (started from the stationary distribution ρ), and we let the process run for a finite time T.

$$P\left(S_\rho(\omega) = T\sigma\right) = \int d\mathcal{P}_\rho(\omega)\delta\left(S_\rho(\omega) - T\sigma\right)$$

where, as before, the integral is taken over all possible paths $\omega = (x_t)_{0\le t\le T}$. Of course, computing this distribution is not possible in general, but we can rewrite it:

$$\begin{aligned} P\left(S_\rho(\omega) = T\sigma\right) &= \int d\mathcal{P}_\rho(\theta\omega)\frac{d\mathcal{P}_\rho}{d\mathcal{P}_\rho\theta}(\omega)\delta\left(S_\rho(\omega) - T\sigma\right) \\ &= e^{T\sigma}\int d\mathcal{P}_\rho(\theta\omega)\delta\left(S_\rho(\omega) - T\sigma\right) \end{aligned}$$

Where in the last line, we used the definition of $S_\rho(\omega)$ and the fact that the delta function makes sure that $S_\rho(\omega) = T\sigma$. We now make a change of variables $\omega \to \theta\omega$, which is an involution. Note that this changes S_ρ to $-S_\rho$:

$$\begin{aligned} P\left(S_\rho(\omega) = T\sigma\right) &= e^{T\sigma} \int d\mathcal{P}_\rho(\omega)\delta\left(S_\rho(\omega) + T\sigma\right) \\ &= e^{T\sigma} P\left(S_\rho(\omega) = -T\sigma\right) \end{aligned}$$

This is what is called a fluctuation theorem for finite times for the quantity S_ρ. However, the local detailed balance assumption only says that $S(\omega)$ is an entropy flux. Without going into mathematical rigour, we do expect that, typically:

$$\lim_{T\to\infty} \frac{S(\omega)}{T} = \lim_{T\to\infty} \frac{S_\rho(\omega)}{T}$$

because the difference between S and S_ρ is only in temporal boundary terms. With this we get the fluctuation theorem (2.21). Note that it is possible that the boundary terms stay important even in the long time limit. In such cases the fluctuation theorem is not valid, see [90].

2.10.2 Work relations

Nonequilibrium work relations [19, 20, 52] can be derived for a system connected to a single heat bath, with a time-dependent dynamics parametrized by a parameter λ_t. This parameter is called the protocol. Of course, because λ_t changes in time, the system is pulled out of equilibrium. Usually it is assumed that for each fixed value λ the dynamics is detailed balanced. Actually, to have work relations, the dynamics only needs to be detailed balanced at the beginning and the end of the process.

We start at time zero with the system prepared in the equilibrium distribution corresponding to the value $\lambda_0 = A$. We then let the process run in the time-dependent dynamics until time T. During the process there is local detailed balance, but at time T there is again a detailed balanced dynamics with $\lambda_T = B$. Consider the following quantity:

$$\frac{d\mathcal{P}_{\rho_A}}{d\mathcal{P}^R_{\rho_B}\theta}(\omega) = \frac{\rho_A(x_0)}{\rho_B(x_T)} e^{S(\omega)}$$

where $S(\omega)$ is by the local detailed balance assumption equal to β times the heat flux $Q(\omega)$ into the environment . Remember that with the time-reversal,

also the protocol λ_t changes to λ_{T-t} (see Section 2.5), hence the superscript R. Substituting the explicit expression of the equilibrium distributions ($\rho_A(x) = \frac{1}{Z_A} e^{-\beta U(A,x)}$ and similarly for $\lambda = B$), we get

$$\frac{d\mathcal{P}_{\rho_A}}{d\mathcal{P}^R_{\rho_B}\theta}(\omega) = e^{\beta[F_A - F_B - U(A,x_0) + U(B,x_T) + Q(\omega)]} = e^{\beta[-\Delta F + W(\omega)]}$$

where $\Delta F = F_B - F_A = \frac{1}{\beta}[\log Z_A - \log Z_B]$ is the change of free energy between the equilibrium states corresponding to the values A and B of the parameter λ, and W is the work done on the system.

From this basic result a work relation can be derived, which was done for the first time by Crooks [19, 20]: let $P(W = w)$ be the probability density that the work during the process is equal to w. Then

$$\begin{aligned}
P(W = w) &= \int d\mathcal{P}_{\rho_A}(\omega)\delta(W(\omega) - w) \\
&= \int d\mathcal{P}^R_{\rho_B}(\theta\omega) e^{\beta[-\Delta F + W(\omega)]} \delta(W(\omega) - w) \\
&= e^{\beta[-\Delta F + w]} \int d\mathcal{P}^R_{\rho_B}(\omega)\delta(W(\omega) + w) \\
&= e^{\beta[-\Delta F + w]} P^R(W = -w)
\end{aligned}$$

where we defined $P^R(W = -w)$ as the probability density that the work during the reversed process is equal to $-w$, and we used that $W(-w) = -W(w)$. From this relation the following equality can easily be derived (obtained by Jarzynski in [52]):

$$\begin{aligned}
\left\langle e^{-\beta W} \right\rangle_{\rho_A} &= \int dw P(W = w) e^{-\beta w} \\
&= \int dw e^{-\beta \Delta F} P^R(W = -w) \\
&= e^{-\beta \Delta F}
\end{aligned}$$

which relates nonequilibrium work-values to equilibrium free energies. Again we see that these results are an immediate consequence of the local detailed balance assumption.

2.10.3 McLennan formula

The McLennan formula [79, 80] gives an approximation of the stationary measure for a dynamics that is close to detailed balance. For a system in contact with two heat baths (particle reservoirs) close to detailed balance means that the difference in the temperatures (chemical potentials) is small, while for a system under the influence of a nonconservative force this means that the force is small. In general, we can parametrize the 'distance' from detailed balance by a parameter ϵ, (e.g. $\epsilon = T_2 - T_1$ or the nonconservative force $F = \epsilon f$). In the case that ϵ is small, one can expand the stationary distribution around $\epsilon = 0$. Up to first order in ϵ this gives the McLennan formula.

We derive the McLennan formula in the following way: suppose that until time $t = 0$ the system is detailed balanced ($\epsilon = 0$), and at time $t = 0$ the system has the corresponding equilibrium distribution $\mu_0 = \rho^0$. At time zero, we drive the system from equilibrium, parametrized by ϵ. After a time t the probability distribution of being in a state x is then given by

$$\mu_t(x) = \langle \delta(x_t - x) \rangle^{\epsilon}_{\rho^0}$$

where the average is an average over all possible paths in the interval $[0, t]$, computed in the dynamics with ϵ (hence the superscript). To rewrite expectation values in the nonequilibrium system into expectation values in the equilibrium system, let us write down the Radon-Nikodym derivative (2.9):

$$e^{-A(\omega)} = \frac{d\mathcal{P}^{\epsilon}_{\rho^0}}{d\mathcal{P}^{0}_{\rho^0}}(\omega)$$

so that the probability distribution becomes:

$$\mu_t(x) = \left\langle \delta(x_t - x) e^{-A(\omega)} \right\rangle^{0}_{\rho^0}$$

with the superscript 0 denoting that the average is taken in the equilibrium process. The expectation value of an observable $Q(\omega)$ in an equilibrium process, is the same as that of the time-reversed observable $Q(\theta\omega)$. This is a simple consequence of (2.14). This means that we have

$$\begin{aligned} \mu_t(x) &= \left\langle \delta(x_0 - x) e^{-A(\theta\omega)} \right\rangle^{0}_{\rho^0} \\ &= \left\langle \delta(x_0 - x) e^{-S_{ex}(\omega)} e^{-A(\omega)} \right\rangle^{0}_{\rho^0} \end{aligned}$$

$$= \left\langle \delta(x_0 - x) e^{-S_{ex}(\omega)} \right\rangle^{\epsilon}_{\rho^0}$$

where the excess entropy flux $S_{ex} = S_\epsilon - S_0$ is the difference in entropy fluxes in the nonequilibrium and the equilibrium system (see (2.19)). One can see from the definition of S_{ex} that it is zero when ϵ is zero. Moreover, in physical systems (and certainly the ones we discuss in this thesis) , S_{ex} is at least of first order in ϵ: $\lim_{\epsilon \to 0} S_{ex}/\epsilon < \infty$. An expansion in ϵ gives

$$\mu_t(x) = \rho^0(x) - \langle \delta(x_0 - x) S_{ex}(\omega) \rangle^0_{\rho^0} + o(\epsilon) = \rho^0(x)[1 - \langle S_{ex}(\omega) \rangle^0_x] + o(\epsilon)$$

where the last expectation is an average over all trajectories starting from the state x. Up to first order in ϵ this formula gives the probability distribution of finding the system in state x at an arbitrary time. Letting the time go to infinity, this will converge to the stationary state $\rho^\epsilon(x)$ corresponding to the nonequilibrium dynamics.

2.10.4 Fluctuation-dissipation theorem in equilibrium

The equilibrium fluctuation-dissipation theorem [13, 59] tells us how systems respond to a small change in their Hamiltonian (energy), at least in the linear regime. The framework in which we work here is almost the same as for the McLennan formula: up to time zero the system enjoys an equilibrium dynamics with a Hamiltonian H_0, and at time zero it has relaxed to the corresponding equilibrium distribution $\rho \propto \exp(-\beta H_0)$. At time zero an extra potential is added to the dynamics: $H = H_0 - hV$, where h is a small parameter. For finite times after this perturbation has been made, the system is not yet in equilibrium. We want to compute the expectation values of observables in this perturbed (nonequilibrium) system. We consider only observables that are state functions, denoting them by $Q(x)$. With a reasoning completely analogous to the one made for the McLennan formula (the only difference is that the observable $\delta(x_t - x)$ is replaced by $Q(x_t)$) one then finds that:

$$\langle Q(x_t) \rangle^h_{\rho^0} = \left\langle Q(x_0) e^{-S_{ex}(\omega)} \right\rangle^h_{\rho^0}$$

where the superscript h denotes that the averages are taken in the perturbed system, and S_{ex} is the excess entropy flux, excess of the perturbed system versus the unperturbed system. As the only difference between the two dynamics is the addition of the potential hV, this excess must be equal to $\beta h[V(x_t) - V(x_0)]$.

In linear perturbation theory one is restricted to systems for which such a small perturbation has only a small influence. One can therefore work up to linear order

in the small parameter h:

$$\begin{aligned}\langle Q(x_t)\rangle^h_{\rho^0} &= \langle Q(x_0)\rangle^0_{\rho^0} - \langle Q(x_0)S_{ex}(\omega)\rangle^0_{\rho^0} + o(h)\\ &= \langle Q(x_0)\rangle^0_{\rho^0} + \langle Q(x_t)S_{ex}(\omega)\rangle^0_{\rho^0} + o(h)\end{aligned}$$

In the last step we have again used that in an equilibrium dynamics the expectations of observables and their time-reversals are the same. Using the explicit expression for the excess entropy flux, we get

$$\frac{\partial}{\partial h}\langle Q(x_t)\rangle^h_{\rho^0}\bigg|_{h=0} = \beta\,\langle Q(x_t)[V(x_t)-V(x_0)]\rangle^0_{\rho^0} \tag{2.22}$$

which is the fluctuation-dissipation theorem in equilibrium. it is valid for any time t. The left hand side of this equation is called a response function. It is the response of the expectation value of an observable to an added potential and is thus equal to the correlation of the observable with the extra entropy flux created by the potential.

The fluctuation-dissipation theorem is useful because it gives a relation between two quantities in essentially different processes. One can e.g. determine the response of a system without actually perturbing it. Part of the results in this thesis are about the generalization of this relation to nonequilibrium systems.

Chapter 3

Markov jump processes

Markov jump processes are Markov processes where the configuration changes discretely. This means that in a finite time-interval, the configuration changes a finite number of times. In this thesis we only work with discrete configuration spaces where Markov jump processes are concerned. Such processes are widely used in physical modelling because they are relatively simple, but approximate real physical systems quite accurately. The discrete configuration space can arise as a result of a discrete approximation of a continuous space, or because of quantum mechanical principles (e.g. spins in a magnetic field). We give a brief introduction to Markov jump processes here, focusing on the properties relevant for the rest of the thesis. For a thorough introduction, we refer to [46]. As a motivation, we start with a very physical example.

3.1 An introductory example

Imagine a chemical reaction where two reactants A and B react to form a product C. The reaction is facilitated by the use of a catalyst X as follows: first the molecule A binds to the catalyst, forming a new molecule X_A. Then B binds, forming X_{AB}. When bound to X, A and B react and form X_C. Finally C decouples from the catalyst, leaving the catalyst in its original state X, see Figure 3.1.

The details of the reaction $A + B \to C$ thus depend strongly on the reaction cycle of the catalyst X. Apart from measuring the average speed or rate of such a cycle, we are also interested in more detailed information. What are the fluctuations from this average speed? Does the inverse cycle also occur, and with what probability? For such questions a more detailed model of the reaction is required. As it is

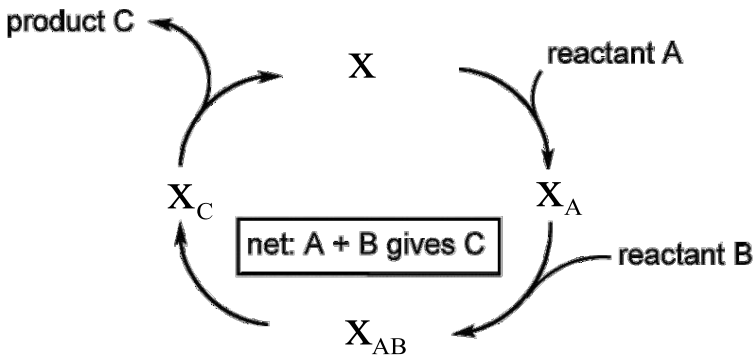


Figure 3.1: A catalytic reaction cycle.

impossible to predict all the precise positions and velocities of all the molecules, a Hamiltonian description is not wanted. Therefore we use a stochastic model. One can imagine all the molecules involved in the reactions are 'wandering' in some solution. A reaction can occur when two molecules meet each other. The speed or rate of each step in the reaction cycle then depends on parameters as concentrations of the molecules in the solvent, temperature, and of course also on the probability that, given that two molecules meet, they actually react. If the concentrations of the molecules in the solvent are not too big, it is a rather good assumption to think of this process as being Markovian. Indeed, we assume that a reactant molecule makes many collisions with the solvent molecules before meeting with another reactant, thus effectively losing its 'memory' on a timescale that is much smaller than we are interested in: the timescale of the successive reactions. Also, this allows us to treat the molecules of the catalyst as independent particles.

To make a model of the reaction cycle, we therefore take one catalyst molecule as the system of interest. This molecule can be in four different chemical states: X, X_A, X_{AB} and X_C. At any time, regardless of its history, the molecule can change its state as a consequence of a reaction. We thus arrive at a Markov process on a finite configuration space, but in continuous time. This fits exactly in the framework of Markov jump processes.

This is only one example where Markov jump processes provide a good model. Other examples are traffic jams, transport of ions through nanotubes, the Ising model and other models of spins in a magnetic field, population dynamics in ecological systems, etc. Providing important and relevant models in physics, Markov jump processes are used throughout this thesis. In this chapter we

therefore introduce the aspects relevant for the rest of this thesis. For a thorough introduction, we refer to [46]. At the end of this chapter we return to this model and see how to describe it as a Markov jump process.

3.2 Definition

We work here with a discrete, even finite, set Ω of configurations $x, y, \ldots$. The dynamics of a Markov jump process is defined as follows: let $P^{\Delta t}(y|x,t)$ denote the probability that the system changes (jumps) from configuration x to $y \neq x$ within $[t, t+\Delta t]$, given that the system was in configuration x at time t. Then

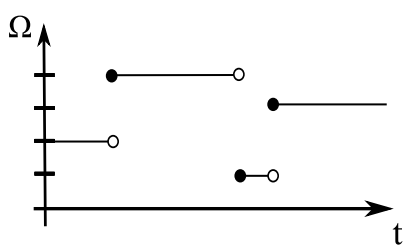
$$P^{\Delta t}(y|x,t) = k_t(x,y)\Delta t + o(\Delta t) \tag{3.1}$$

The $k_t(x,y)$ are called transition rates. Obviously $k_t(x,y) \geq 0$. Also, as a convention, we take $k_t(x,x) = 0$. From (3.1) one also sees that the probability to jump two times within a time span Δt is of order $o(\Delta t)$. Logically, the probability $P^{\Delta t}(x|x,t)$ that the system does not jump within $[t, t+\Delta t]$ is then

$$P^{\Delta t}(x|x,t) = 1 - \lambda_t(x)\Delta t + o(\Delta t) \quad \text{with} \quad \lambda_t(x) = \sum_{y \neq x} k_t(x,y) \tag{3.2}$$

This $\lambda_t(x)$ is called the escape rate, as it quantifies the probability that the system 'escapes' from x. We assume that $\epsilon \leq \lambda_t(x) < \lambda$ for all t, x and for some $\epsilon > 0, \lambda < \infty$, to make sure that the expected time for the system to jump is finite.

In Figure 3.2, an example of a typical realization of a Markov jump process is shown: the system stays in a certain configuration for an exponentially distributed time (we prove this later on in this chapter), and then jumps to the next state. We adopt the convention that at the jump time, the system is in the state after



Figuur 3.2: A realization of a Markov jump process.

the jump, i.e. in Figure 3.2 the graph is right-continuous.

3.3 Probability distributions and the Master equation

Suppose that we know the probability that the system is in state x at time t: $\mu_t(x)$, for all x. From this we can compute the probability that the system is in any state x at time $t + \Delta t$. It is equal to the probability that the system was in state x at time t and stayed there, plus the probability that it was in any state y and jumped to x. With (3.1) and (3.2) it is clear that

$$\mu_{t+\Delta t}(x) = \mu_t(x)[1 - \lambda_t(x)\Delta t + o(\Delta t)] + \sum_{y \neq x} \mu_t(y)[k_t(y,x)\Delta t + o(\Delta t)]$$

This gives us a differential equation for μ_t:

$$\frac{d\mu_t(x)}{dt} = \lim_{\Delta t \to 0} \frac{\mu_{t+\Delta t}(x) - \mu_t(x)}{\Delta t} = \sum_{y \neq x} [\mu_t(y)k_t(y,x) - \mu_t(x)k_t(x,y)] \quad (3.3)$$

This equation is called the Master equation. Notice that it is a deterministic equation: all uncertainty is in the probability distribution μ_t, while the evolution of μ_t through time is deterministic.

The quantity $j_{\mu_t}(x,y) = \mu_t(x)k_t(x,y) - \mu_t(y)k_t(y,x)$ can be seen as the current between x and y, or better even: a probability current. Indeed, the first term is the probability per unit of time that the system makes a jump from x to y, while the second term is the probability per unit of time that the system makes a jump from y to x. With this definition of current we can rephrase the Master equation as a continuity equation:

$$\frac{d\mu_t(x)}{dt} + \sum_{y \neq x} j_t(x,y) = 0 \quad (3.4)$$

3.4 Stationarity and detailed balance

In the special case that the transition rates of the Markov process are independent of time ($k_t(x,y) = k(x,y)$), we call the process time-homogeneous. As said before, for time-independent dynamics we assume that there exists a stationary

distribution ρ, which is unique, and that all distributions μ_t converge to it in the long time limit. This ρ solves the Master Equation (3.3) with the left-hand side zero. The conditions under which this assumption is valid are conditions of irreducibility: a jump process is called irreducible if the probability of the system to go from any state x to any other state y in a finite time is non-zero. This is equivalent to saying there is a chain of states $x = x_1, x_2, \ldots, x_n = y$ such that $k(x_i, x_{i+1}) > 0$ for all $i = 1$ to n. The proof of this is beyond the scope of this text.

Finding the stationary distribution is difficult in general. There is however a special case in which computations simplify: if there exists a function $f(x)$, ($f \neq 0$ and $f(x) \geq 0$), such that

$$f(x)k(x,y) = f(y)k(y,x) \tag{3.5}$$

then $\rho(x) \propto f(x)$ solves the stationary Master Equation. If the transition rates satisfy this property, then the process is said to satisfy detailed balance, and ρ is called the equilibrium distribution. Once the system has relaxed to this distribution, we then also have that $j_\rho(x,y) = 0$, so that in equilibrium all currents between different states are zero. This definition of equilibrium coincides with the one given in the previous chapter.

3.5 Finite time probabilities

To have a better idea of how the system evolves from configuration to configuration, we compute the following probability: for any $t > s$, let $P(y,t|x,s)dt$ be the probability that the system remains in configuration x until it jumps within $[t, t+dt]$ to state y, given that the system was in configuration x at time s. Then

$$P(y,t|x,s)dt = k_t(x,y)e^{-\int_s^t \lambda_u(x)du}dt \tag{3.6}$$

We prove this as follows: divide the interval $[s,t]$ into n pieces of length $\Delta t = \frac{t-s}{n}$. Then from (3.1) and (3.2) we see that

$$\begin{aligned} P(y,t|x,s)dt &= k_t(x,y)dt \prod_{k=0}^{n-1} [1 - \lambda_{s+k\Delta t}(x)\Delta t + o(\Delta t)] \\ &= \exp\left\{ \sum_{k=0}^{n-1} \ln\left[1 - \lambda_{s+k\Delta t}(x)\Delta t + o(\Delta t)\right] \right\} k_t(x,y)dt \end{aligned}$$

$$= \quad \exp\left\{\sum_{k=0}^{n-1}\left[-\lambda_{s+k\Delta t}(x)\Delta t + o(\Delta t)\right]\right\} k_t(x,y)dt$$

If we now let $n \to \infty$, we get exactly (3.6).

Let us examine this probability: the probability to jump to anywhere within $[t, t+dt]$, given that the system was in state x at time s is

$$P(t|x,s)dt = \sum_y P(y,t|x,s)dt = \lambda_t(x)e^{-\int_s^t \lambda_u(x)du}dt$$

On the other hand, given that the system jumps at time t, the probability that it jumps to state y is then

$$p_t(x,y) = \frac{P(y,t|x,s)dt}{P(t|x,s)dt} = \frac{k_t(x,y)}{\lambda_t(x)}$$

The $p_t(x,y)$ are called the transition probabilities, and we see that $\sum_{y\neq x} p_t(x,y) = 1$. In conclusion, when at time s the system is in a state x, it stays there for a time $t-s$ that is exponentially distributed, determined by the escape rate $\lambda_t(x)$. When it jumps, the probability that it jumps to a state y is $p_t(x,y)$.

3.6 Trajectories

A trajectory $\omega = (x_t)_{0\leq t\leq T}$ is completely described by giving the consecutive states x_i (with $i = 0, \ldots n$) the system visits plus the times t_i at which it jumped:

$$x_t = x_i \quad \text{for} \quad t_i \leq t < t_{i+1}$$

where $t_0 = 0$ and t_n is the last jump time before time T. Using (3.6), we can compute the probability measure of such a path:

$$d\mathcal{P}_{\mu_0}(\omega) = \mu_0(x_0)\prod_{i=0}^{n-1}\left[k_{t_{i+1}}(x_i,x_{i+1})e^{-\int_{t_i}^{t_{i+1}}\lambda_u(x_i)du}dt_{i+1}\right]e^{-\int_{t_n}^{T}\lambda_u(x_n)du}$$

The last factor on the right-hand side is the probability that the system does not jump in the interval $[t_n, T]$. We rewrite this path-probability measure in a slightly more elegant form:

$$d\mathcal{P}_{\mu_0}(\omega) = \mu_0(x_0) \exp\left\{ \sum_{i=0}^{n-1} \log k_{t_{i+1}}(x_i, x_{i+1}) - \int_0^T \lambda_t(x_t) dt \right\} dt_1 \dots dt_n \tag{3.7}$$

With this path-probability one can then compute expectation values, as defined in (2.2), and the time-evolved probability distribution of configurations μ_t, as defined in (2.5). This definition of μ_t is completely consistent with the Master equation, of course.

3.7 The Girsanov formula

Suppose now that one has a different Markov jump process with rates $k^*(x, y)$ and one wants to calculate the expectation values of observables in this process. Then one can relate this to the expectation values in the original process by using the Radon-Nikodym derivative (see (2.8) and (2.9)). The demand of absolute continuity boils down to the demand that $k^*(x, y) \neq 0$ for any pair x, y for which $k(x, y) \neq 0$. It follows immediately from (3.7) that

$$e^{-A(\omega)} = \exp\left\{ \sum_{i=0}^{n-1} \ln\left(\frac{k^*_{t_{i+1}}(x_i, x_{i+1})}{k_{t_{i+1}}(x_i, x_{i+1})} \right) + \int_0^T dt[\lambda_t(x_t) - \lambda^*_t(x_t)] \right\} \tag{3.8}$$

with the action defined as in (2.9) This is often schematically written as

$$A(\omega) = \int_0^T dt[\lambda^*_t(x_t) - \lambda_t(x_t)] - \sum_{t \leq T} \ln\left(\frac{k^*_t(x_{t^-}, x_t)}{k_t(x_{t^-}, x_t)} \right) \tag{3.9}$$

where the sum is over jump times, and t^- is the time just before the jump. This formula (3.9), the Radon-Nikodym derivative for Markov jump processes, is called the Girsanov formula, and is of much use throughout this text. More details can be found in Appendix 2 of [57].

3.8 Local detailed balance

The Markov jump processes as described up to now are little more than mathematical models. To make a connection with physics we make the local

detailed balance assumption: the quantity $S(\omega)$ that was defined in (2.12), is equal to the entropy flux into the environment during the trajectory ω. For Markov jump processes this becomes (see (3.7)):

$$S(\omega) = \log\left(\frac{d\mathcal{P}_{x_0}}{d\mathcal{P}^R_{x_T}\theta}(\omega)\right) = \sum_{t\leq T} \log\frac{k_t(x_{t^-}, x_t)}{k_t(x_t, x_{t^-})}$$

We see that there is only an entropy flux when there is a jump, and the total entropy flux is just the sum of entropy fluxes for each jump:

$$\frac{k_t(x,y)}{k_t(y,x)} = e^{\sigma_t(x,y)}$$

where $\sigma_t(x,y)$ is the entropy flux associated to one jump from a state x to a state y.

Example 1: As a first example we imagine a system with time-homogeneous dynamics, immersed in a single heat bath at inverse temperature β. We denote the energy of the system, (which depends on the configuration x) as $U(x)$. On top of that there is a forcing, meaning that for every jump from a state x to a state y, there is an amount of work $W(x,y)$ needed. Note that by definition $W(x,y) = -W(y,x)$. Using the first law of thermodynamics, we see that

$$\frac{k(x,y)}{k(y,x)} = e^{-\beta[U(y)-U(x)-W(x,y)]} \tag{3.10}$$

Indeed, local detailed balance allows us to express a part of the transition rates in terms of physical, measurable quantities.

If there exists a state function V such that for all x, y we have $W(x,y) = V(y) - V(x)$, then the forcing is conservative, and by (3.10) the rates are detailed balanced with the corresponding equilibrium distribution $\rho(x) = \frac{1}{Z}\exp\{-\beta[U(x) - V(x)]\}$.

Example 2: As a second example think of two big particle reservoirs with different particle densities. Between these reservoirs there is a wall, with a narrow channel going through it from one side to the other. The channel is so narrow that particles are effectively restricted to one dimension. This is what happens for example in ionic transport through cell membranes. Because of the difference in particle densities a current arises towards the side with the lower density.

We model the channel as a set of sites labelled by $i = 1, \ldots, N$. On each site there can either be one or no particle. A configuration of the system is thus an

array x, with empty sites $x_i = 0$ and particles $x_i = 1$. Each particle can hop to the neighbouring site left or right if that site is unoccupied. Such a process is called an exclusion process. Hence the only transitions allowed are $x \to x^{(ij)}$ where $j = i \pm 1$ and

$$x_k^{(ij)} = \begin{cases} x_k & \text{if} \quad i \neq k \neq j \\ x_j & \text{if} \quad k = i \\ x_i & \text{if} \quad k = j \end{cases} \tag{3.11}$$

An exception occurs at the boundaries of the system $i = 1, N$, where particles can enter and leave the system. Hence transitions $x \to x^{(i)}$ are also allowed, with $i = 1, N$ and

$$x_k^{(i)} = \begin{cases} x_k & \text{if} \quad k \neq i \\ 1 - x_k & \text{if} \quad k = i \end{cases} \tag{3.12}$$

This means that at the left and right boundary the system is in contact with particle reservoirs, which are characterized by a temperature and a chemical potential. In Figure 3.3 an example of such a process is visualized. This is an example of Kawasaki dynamics.

Physically, we assume here that there is no interaction between the particles, except for the exclusion. We take therefore the energy of the system zero. The particle number of this system is not conserved however. If both reservoirs have the same inverse temperature β and chemical potential μ, the system is detailed balanced, and for any two states x, y:

$$\frac{k(x,y)}{k(y,x)} = e^{\beta\mu[n(y)-n(x)]}$$

with $n(x)$ the number of particles in configuration x. We see from this that when a particle hops within the system, the transition rate is symmetric: $k(x, x^{(ij)}) = k(x^{(ij)}, x)$. Detailed balance gives us as equilibrium distribution

$$\rho(x) = \frac{e^{\beta\mu n(x)}}{(1 + e^{\beta\mu})^N}$$

and the average particle density d

$$d = \sum_x \rho(x) \frac{n(x)}{N} = \frac{e^{\beta\mu}}{1 + e^{\beta\mu}}$$

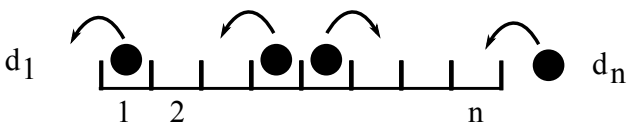
This d is the particle density of the particle reservoir, as it is equal to the density of the system when in equilibrium.

If the chemical potentials of the two reservoirs are not the same, we use local detailed balance. We get our inspiration from the detailed balance case to write

$$\frac{k(x, x^{(i)})}{k(x^{(i)}, x)} = e^{\beta\mu_i[n(x^{(i)})-n(x)]} \tag{3.13}$$

for events where a particle enters or leaves the system ($i = 1, N$), and $k(x, x^{(ij)}) = k(x^{(ij)}, x)$ for events where a particle hops within the system. Again inspired by the detailed balance case we define the particle densities of the reservoirs as

$$d_i = \frac{e^{\beta\mu_i}}{1 + e^{\beta\mu_i}}$$



Figuur 3.3: A visualization of an exclusion process.

Finding the stationary distribution in this case is possible, but certainly not trivial. We refer to [25] for more details. We can say something about the current though. For this we use the fluctuation theorem discussed in Section 2.10.1. For any specific ω we can write, using (3.13)

$$S(\omega) = \beta T\mu_1 J_1(\omega) + \beta T\mu_N J_N(\omega)$$

where the current $J_i(\omega)$ is the net number of particles per unit of time that has gone into the system through site i. Note that the sum of these two currents gives the change of particle number in the system per unit of time: $J_1(\omega) + J_N(\omega) = [n(x_T) - n(x_0)]/T$. Therefore:

$$\lim_{T\to\infty}[J_1(\omega) + J_N(\omega)] = 0$$

In the limit of large T we therefore just use $J = J_1 = -J_N$. The fluctuation theorem (2.21) can then be reformulated as

$$\lim_{T\to\infty}\frac{1}{T}\log\left(\frac{P\left(J(\omega) = j\right)}{P\left(J(\omega) = -j\right)}\right) = \beta j(\mu_1 - \mu_N)$$

The local detailed balance assumption on which the fluctuation theorem is based thus predicts what is the most probable direction of the particle current. Not surprisingly the most probable direction of the current is towards the lower chemical potential. The fluctuation theorem asserts that the reverse current is exponentially less probable.

Example 3: Let us return to the introductory example from the beginning of this chapter. More precisely: Ω consists of 4 chemical states , labelled by $x = 1, \ldots, 4$ and lying on a ring, i.e. $5 \equiv 1$. These states correspond to: $1 \equiv X$, $2 \equiv X_A$, $3 \equiv X_{AB}$ and $4 \equiv X_C$. We can see that $k(x, y)$ only differs from zero for $|y-x| = 1$. Let us denote by $U(x)$ the energy of the molecule when it is in chemical state x, and μ_i and N_i are the chemical potentials and particle numbers of the reactant particles in the solvent, where $i = A, B, C$. These reactant particles are seen as the reservoirs of our catalyst molecule. As we imagine the reservoirs to be in equilibrium, their chemical potentials are fixed. The entropy change in the environment for a transition $x \to y$ can thus be written as

$$\sigma(x, y) = -\beta[U(y) - U(x)] - \beta \sum_i \mu_i \Delta N_i$$

where ΔN_i is the change in particles of species i. For example, if the catalyst molecule makes a transition $X \to X_A$, we get as entropy flux to the environment:

$$\log \frac{k(1, 2)}{k(2, 1)} = \sigma(1, 2) = -\beta[U(2) - U(1)] + \beta\mu_A$$

It is the chemical potentials of the 'reservoirs' that drive the system from equilibrium. To show this, we take as trajectory one reaction cycle $X \to X_A \to X_{AB} \to X_C \to X$. We can compute the entropy flux into the environment during this cycle:

$$\sum_{i=1}^{4} \sigma(i, i + 1) = \beta[\mu_A + \mu_B - \mu_C]$$

For an equilibrium dynamics this entropy flux should be equal to zero (see (2.15)). Whenever $\mu_C \neq \mu_A + \mu_B$, the probability that the catalyst covers the reaction cycle in one direction is not equal to the probability for the other direction. In fact, when $\mu_C < \mu_A + \mu_B$ the reaction cycle will on average go in the direction as in Fig. 3.1.

Chapter 4

Diffusions

Langevin dynamics (diffusions) are well-known and widely used as a combination of Hamiltonian dynamics and stochastic noise. Mathematically they are more difficult to define than Markov jump processes, but their physical meaning is much more clear, as the dynamics are expressed in terms of measurable quantities as forces and positions, or electric fields and currents. One distinguishes in diffusions between the underdamped (inertial) regime and the overdamped regime. In the latter accelerations are ignored due to high friction. In this chapter we briefly introduce diffusions, treating both the underdamped and the overdamped case. For the interested reader, we suggest [18, 58, 78, 81].

4.1 The Langevin equation

For didactic purposes we start with the simplest case of one test-particle in one dimension. The state space is here the phase space, meaning the space of the possible positions and velocities of the particle. Pure Hamiltonian dynamics would give two equations of motion:

$$\begin{aligned} dx_t &= v_t dt \\ m dv_t &= F_t(x_t) dt \end{aligned} \tag{4.1}$$

which correspond to a particle that experiences a deterministic force F_t. Note that m is the mass of the particle, x_t its position at time t and v_t its velocity. Now we put the particle in a heat bath (imagine a glass of water) at inverse temperature β. The particles of the heat bath will start colliding with our test-particle. Because

we do not know all the exact movements of these particles we model the effect of these collisions as a random force on the test-particle. The modified equations then look like this:

$$\begin{aligned} dx_t &= v_t dt \\ m dv_t &= F_t(x_t)dt - m\gamma v_t dt + \sqrt{2\mathcal{D}}dB_t \end{aligned} \tag{4.2}$$

and are called the Langevin equations. The term $-m\gamma v_t$ is the frictional force that the test-particle feels, with γ the friction coefficient. Indeed, when the particle has a velocity, then it has more head-on collisions than from behind, with a net effect equal to $-m\gamma v_t$. This friction force is an average effect of the heat bath. That is why still another term is added, namely $\sqrt{2\mathcal{D}}dB_t$, with $\mathcal{D}$ the diffusion coefficient. This should model the random kicks the particle receives. B_t is what is called a Wiener process. It is a random variable with the property $B_t - B_s \sim \mathcal{N}(0, t-s)$, meaning that it has independent increments with a Gaussian distribution with mean zero and variance $t - s$. This is the reason that dB_t is usually called a standard Gaussian white noise. The random variable $\xi_t = \frac{dB_t}{dt}$ is often used here, but mathematically ill-defined. It also has mean zero, but has the property that $\langle \xi_s \xi_t \rangle = \delta(t-s)$. This means that we have approximated the physical process of collisions with particles of the bath, as a succession of infinitely many, infinitely big kicks. As a consequence, dv_t is not really well-defined. Still the equations (4.2) are mathematically well-defined if one integrates them with respect to time. This is not done here because the Langevin equations do have a clear physical meaning in this form (4.2).

An important remark for things to follow is that we interpret dv_t in the Langevin equation as $v_{t+dt} - v_t$, and dx_t analogously. This is called the Itô-interpretation. It is also true that the inverse temperature β of the heat bath has not yet entered the description. We come back to this at the end of this chapter, when we discuss the local detailed balance assumption.

4.2 The Fokker-Planck equation and stationarity

We sample the initial positions and velocities (at time zero) from an initial probability density $\mu_0(x, v)$. One can then derive (see Appendix B) from the Langevin equation an evolution equation for the densities at later times, which is called the Fokker-Planck equation:

$$\frac{\partial \mu_t}{\partial t} = -v\frac{\partial \mu_t}{\partial x} - \frac{\partial}{\partial v}\left[(\frac{F_t}{m} - \gamma v)\mu_t\right] + \frac{\mathcal{D}}{m^2}\frac{\partial^2 \mu_t}{\partial v^2} \tag{4.3}$$

Just as in the case of Markov jump processes, this evolution equation is deterministic, and we can rewrite it in the form of a continuity equation:

$$\frac{\partial \mu_t}{\partial t} + \nabla \cdot j_{\mu_t} = 0 \tag{4.4}$$

with a probability current defined by

$$j_{\mu_t} = \left(v\mu_t, (\frac{F_t}{m} - \gamma v)\mu_t - \frac{\mathcal{D}}{m^2}\frac{\partial \mu_t}{\partial v} \right) \tag{4.5}$$

When the force is time-independent, we assume as always that there is a stationary distribution ρ, (which solves the Fokker-Planck equation with left-hand side zero). Unfortunately, it is often impossible to just solve the Fokker-Planck equation.

In the special case however that we have $F(x) = -\frac{\partial U(x)}{\partial x}$, then the solution is immediately found to be

$$\rho(x, v) = \frac{1}{Z} e^{-\beta[U(x) + \frac{mv^2}{2}]}$$

where $\beta = \frac{m\gamma}{\mathcal{D}}$. Indeed, in this special case the particle is in equilibrium with the surrounding fluid at inverse temperature β and has the well-known Gibbs distribution, giving the probability of a state (x, v) as the exponential of minus β times the energy of the state. There is one difference with the Markov jump case, namely that the probability current in equilibrium is not zero here:

$$j_\rho = \left(v\rho, -\frac{1}{m}\frac{\partial U}{\partial x}\rho \right)$$

On the other hand, this is equal to the current one would have in a purely Hamiltonian (deterministic) dynamics, so one could say that in equilibrium, the diffusive part of the current is zero.

4.3 Trajectories

The position and velocity of the particle evolve according to the Langevin equation, and thus follow a trajectory ω through phase-space: $\omega = (x_t, v_t)_{0 \le t \le T}$. One can compute the probability of such a path. We give a heuristic derivation of this, for a mathematically rigorous treatment one should consult [65].

Given the initial position and velocity of the particle, the trajectory ω is completely determined by the increments of the Wiener process. Let us consider the case that x_t and v_t are given, and compute the probability of having (x_{t+dt}, v_{t+dt}) an infinitesimal time later. We know from the Langevin equation (4.2) that

$$dB_t = \frac{1}{\sqrt{2\mathcal{D}}}[mdv_t - F_t(x_t)dt + m\gamma v_t dt]$$

And the fact that $dB_t = B_{t+dt} - B_t \sim \mathcal{N}(0, dt)$ means that we have a probability density P given by:

$$\begin{aligned} dP(v_{t+dt}, x_{t+dt}|x_t, v_t) &= P(v_{t+dt}, x_{t+dt}|x_t, v_t)dx_t dv_t \\ &= \frac{1}{N}\delta(x_{t+dt} - x_t - v_t dt)e^{-L(x_t, v_t, dv_t)}dx_t dv_t \quad (4.6) \end{aligned}$$

where L is given by

$$\begin{aligned} L(x_t, v_t, dv_t) &= \frac{1}{4\mathcal{D}dt}[mdv_t - F_t(x_t)dt + m\gamma v_t dt]^2 \\ &= \frac{dt}{4\mathcal{D}}[m\frac{dv_t}{dt} - F_t(x_t) + m\gamma v_t]^2 \end{aligned}$$

Note that $N = \frac{\sqrt{4\pi\mathcal{D}dt}}{m}$ is a normalization constant. Furthermore we know that $x_{t+dt} = x_t + v_t dt$, which is why a delta function is inserted. From now on, we restrict the space of all trajectories ω to those trajectories which satisfy this constraint. This means that the delta function is no longer necessary. Let us now divide the interval $[0, T]$ into n intervals of length Δt (so $n\Delta t = T$). We should take Δt small enough, so that (4.6) can be used. (We take the limit of $\Delta t \to 0$). Define $t_i = i\Delta t$ for $i = 0, \ldots, n$, then the probability density of ω is just the product of the probability densities of all the steps $v_{t_i} \to v_{t_{i+1}}$, because the increments of a Wiener process are mutually independent:

$$\mathcal{P}_{\mu_0}(\omega) = \frac{\mu_0(x_0, v_0)}{N^n}\exp\left\{-\frac{1}{4D}\sum_{i=0}^{n-1}\Delta t[m\frac{v_{t_{i+1}} - v_{t_i}}{\Delta t} - F_t(x_{t_i}) + m\gamma v_{t_i}]^2\right\}$$

which we denote in the limit of $\Delta t \to 0$ by

$$\mathcal{P}_{\mu_0}(\omega) = \frac{\mu_0(x_0, v_0)}{\mathcal{N}}\exp\left\{-\frac{1}{4\mathcal{D}}\int_0^T dt[m\dot{v}_t - F_t(x_t) + m\gamma v_t]^2\right\}$$

where $\mathcal{N}$ is a normalization constant, and $\mu_0(x_0, v_0)$ is the probability distribution from which we sampled x_0 and v_0. The integral and $\dot{v}_t$ in the exponent are just notation, because they are not really well-defined (like 'Riemann integrals' and 'derivatives') as a consequence of the properties of the Wiener process. Again, we do not go into the full mathematical details here (for that, see [65]). A more important quantity for the purpose of this thesis arises if we compare the path-probabilities of different dynamics to each other, by considering the action of the process with respect to a reference process. For this reference we define a new Langevin dynamics:

$$\begin{aligned} dx_t &= v_t dt \\ m dv_t &= -m\gamma v_t dt + \sqrt{2\mathcal{D}} dB_t \end{aligned} \tag{4.7}$$

Note that it is the same as before, only we have taken the force F to be zero. We denote the path-probability measure of this process by $d\mathcal{P}^0$. Then we have that

$$e^{-A(\omega)} = \frac{d\mathcal{P}_{\mu_0}}{d\mathcal{P}^0_{\mu_0}}(\omega) = \exp\left\{ \frac{1}{4\mathcal{D}} \int_0^T dt[2m\dot{v}_t + 2m\gamma v_t - F_t(x_t)]F_t(x_t) \right\} \tag{4.8}$$

This is the Girsanov-formula for our one-dimensional diffusion process. This quantity is well-defined, as the term in the exponent with $\dot{v}_t$, actually a notation for

$$\int_0^T dt \dot{v}_t F_t(x_t) = \lim_{n\to\infty} \sum_{i=0}^{n-1} (v_{t_{i+1}} - v_{t_i}) F_{t_i}(x_{t_i}) \tag{4.9}$$

is mathematically well-defined. However, because of the properties of the Wiener process we can't treat this 'integral' with the normal (Riemann) rules of integration, and have entered the domain of stochastic integrals. The integral as defined in (4.9) is called a 'stochastic integral in the Itô interpretation,' or just an 'Itô integral.' We explain more on these stochastic integrals at the end of this section. We conclude for now by saying that, like in the previous chapters, we use this path-probability density to define expectation values for observables. If one defines the time-evolved probability density as $\mu_t(x, v) = \langle \delta(x_t - x)\delta(v_t - v) \rangle_{\mu_0}$, then one can prove that its evolution equation is exactly the Fokker-Planck equation, see Appendix B.

4.4 Overdamped diffusions

In many physical cases (e.g. molecular motors) the system is submerged in a highly viscous environment. This means that the friction coefficient γ is very big. If on

top of that the mass of the system is very small, one enters the overdamped region, in which the system after each random kick relaxes quickly to the local expected velocity, determined by the force F and the friction. Because of this, one is not so much interested in the changes of the velocity anymore, just in the positions.

Mathematically one takes the limit of $\gamma \to \infty$ and $m \to 0$, while keeping $\chi = \frac{1}{m\gamma}$ constant and finite. This limit is not trivial [94, 99, 104], but the result is that one can ignore the term mdv_t in the Langevin equation (4.2). If one then replaces in (4.2) the term $m\gamma v_t dt$ by $m\gamma dx_t$, one is left with the single equation

$$dx_t = \chi F_t(x_t)dt + \sqrt{2D}dB_t \tag{4.10}$$

with $D = \chi^2 \mathcal{D}$. We call this the overdamped Langevin equation. To distinguish the previous Langevin equation from this one, we call from now on (4.2) the underdamped Langevin equation. The corresponding Fokker-Planck equation for probability distributions is

$$\frac{\partial \mu_t}{\partial t} + \frac{\partial}{\partial x} j_{\mu_t} = 0 \quad \text{with} \qquad j_{\mu_t} = \chi F_t \mu_t - D\frac{\partial \mu_t}{\partial x} \tag{4.11}$$

which defines the probability current. This equation is in the literature often called the Smoluchowski equation, and the high-friction limit the Smoluchowski limit. In the special case that $F(x) = -\frac{\partial U(x)}{\partial x}$, we can easily find the stationary distribution ρ:

$$\rho(x) = \frac{1}{Z} e^{-\beta U(x)}$$

with $\beta = \frac{\chi}{D}$. Just as in the underdamped case, this corresponds to equilibrium. Moreover, in this case the stationary probability current j_ρ is actually zero.

Let us also provide a simple example of a nonequilibrium dynamics: suppose the particle moves on a circle of length 1. We take the force $F(x) = f$ a constant. This can't be written as the derivative of a potential. One can solve the stationary Fokker-Planck equation, though: the stationary distribution is just the uniform distribution $\rho(x) = 1$. The stationary current is $j_\rho = \chi f$, which is a constant but is not zero.

4.5 Overdamped diffusions in more dimensions

To be more general, let us work in a d-dimensional space, and let x_t be a d-dimensional vector, representing the position of a single particle. Then the most general set of d overdamped Langevin equations, one for each component of x_t, is

$$dx_t^i = \sum_j [\chi_{ij}(x_t)F_t^j(x_t)dt + \frac{\partial D_{ij}}{\partial x^j}(x_t)dt + \sqrt{2D(x_t)}_{ij}dB_t^j] \tag{4.12}$$

In this equation F_t still represents a force, D and χ have become $d \times d$ symmetric matrices both depending on the position x_t, and there are d Wiener processes dB_t^j. The term with the derivatives of D is a result of the fact that a non-constant diffusion coefficient gives an extra net force on the particle [101]. We also stress that we make no difference between indices that are placed as superscripts or as subscripts, i.e. $x^i = x_i$. To reduce notation, we schematically write these Langevin equations in a matrix-form:

$$dx_t = \chi(x_t)F(x_t)dt + \nabla \cdot D(x_t)dt + \sqrt{2D(x_t)}dB_t \tag{4.13}$$

In such a notation, χF is to be interpreted as the matrix χ acting on the vector F. The corresponding Fokker-Planck equation is then

$$\frac{\partial \mu_t}{\partial t} + \nabla \cdot j_{\mu_t} = 0 \quad \text{with} \quad j_{\mu_t} = \chi F_t \mu_t - D\nabla \mu_t \tag{4.14}$$

The path-probability density is found in the same way as for underdamped diffusions:

$$\mathcal{P}_{\mu_0}(\omega) = \frac{\mu_0(x_0)}{\mathcal{N}} \exp\left\{ -\frac{1}{4}\int_0^T dt[\dot{x}_t - \chi F_t - \nabla \cdot D] \cdot D^{-1}[\dot{x}_t - \chi F_t - \nabla \cdot D] \right\}$$

where a dot $\cdot$ denotes a scalar product of vectors. Again it is mathematically safer to look at the Radon-Nikodym derivative of this process with respect to a reference. Let us take as a reference (4.13) with the force F put to zero. We denote the corresponding path-probability by $\mathcal{P}^0$ and get

$$e^{-A(\omega)} = \frac{d\mathcal{P}_{\mu_0}}{d\mathcal{P}^0_{\mu_0}}(\omega) = \exp\left\{ \frac{1}{4}\int_0^T dt[2\dot{x}_t - 2\nabla \cdot D - \chi F_t] \cdot D^{-1}\chi F_t \right\} \tag{4.15}$$

This formula is the Girsanov-formula for overdamped diffusions where, like in the underdamped case, we still have stochastic integrals in our formula, meaning integrals which contain $\dot{x}$.

4.6 Stochastic integrals

In this section we shortly review some technical aspects of stochastic integration. For more information, see [43, 58].

Itô integral Let us denote by q the configuration of the system. For underdamped diffusions this is (x, v), and for overdamped diffusions this is x. In (4.9) and (4.15) one encounters integrals of the form $\int_0^T dt \dot{q}_t^k g_t(q_t)$, usually written as $\int_0^T dq_t^k g_t(q_t)$, with g_t some function of the configuration and q^k an arbitrary component of q. This is actually just notation for the following: divide the interval $[0, T]$ into n intervals of length Δt (so $n\Delta t = T$) and define $t_i = i\Delta t$ for $i = 0, \ldots n$, then for any function $g_t(q)$

$$\int_0^T dq_t^k g_t(q_t) = \lim_{n\to\infty} \sum_{i=0}^{n-1} (q_{t_{i+1}}^k - q_{t_i}^k) g_t(q_{t_i}) \tag{4.16}$$

Because of the special properties of Wiener processes, an integral defined as in (4.16) can not be treated as a Riemann integral. For example, for a Riemann integral, it does not matter where the function g_t is evaluated in each time interval, for stochastic integrals like (4.16) it does. From the way we defined the Langevin equation it is natural that g_t is evaluated at the beginning of each interval. A stochastic integral with this property is called an Itô integral.

To compute what the integration rules are for Itô integrals, note the following: the increments of the Wiener process $B_{t+\Delta t} - B_t$ have a Gaussian distribution with zero mean and variance Δt. From this, one can prove that the quantity $(B_{t+\Delta t} - B_t)^2$ has a mean Δt and a variance of order $(\Delta t)^2$. Therefore, one can safely write $(B_{t+\Delta t} - B_t)^2 = \Delta t + o(\Delta t)$. Analogously: $(q_{t_{i+1}}^k - q_{t_i}^k)(q_{t_{i+1}}^j - q_{t_i}^j) = 2D_{kj}\Delta t + o(\Delta t)$, for any $k, j = 1, \ldots, d$ (Using $\mathcal{D}$ instead of D for underdamped diffusions).

We compute here what the actual rules for integration are for Itô integrals. Take an arbitrary function $g(q)$ not explicitly dependent on time. A Taylor-expansion gives:

$$\begin{aligned} g(q_{t_{i+1}}) &= g(q_{t_i}) + \sum_j (q_{t_{i+1}}^j - q_{t_i}^j) \cdot \frac{\partial}{\partial q^j} g(q_{t_i}) \\ &\quad + \frac{1}{2} \sum_{j,k} (q_{t_{i+1}}^j - q_{t_i}^j)(q_{t_{i+1}}^k - q_{t_i}^k) \frac{\partial^2}{\partial q^j \partial q^k} g(q_{t_i}) + o(\Delta t) \end{aligned}$$

$$\approx \quad g(q_{t_i}) + (q_{t_{i+1}} - q_{t_i}) \cdot \nabla g(q_{t_i}) + (D\nabla) \cdot \nabla g(q_{t_i}) \Delta t + o(\Delta t)$$

As a consequence

$$\begin{aligned} \int_0^T dq_t \cdot \nabla g(q_t) &= \lim_{n\to\infty} \sum_{i=0}^{n-1} (q_{t_{i+1}} - q_{t_i}) \cdot \nabla g_t(q_{t_i}) \\ &= \lim_{n\to\infty} \sum_{i=0}^{n-1} [g(q_{t_{i+1}}) - g(q_{t_i}) - (D\nabla) \cdot \nabla g(q_{t_i}) \Delta t + o(\Delta t)] \\ &= g(q_T) - g(q_0) - \int_0^T dt (D\nabla) \cdot \nabla g(q_t) \end{aligned}$$

The integral over time on the right-hand side is a normal Riemann integral. Similarly one proves that for explicitly time-dependent functions $g_t(q)$ one gets

$$\int_0^T [dq_t \cdot \nabla g_t(q_t) + dt \frac{\partial g}{\partial t}(q_t)] = g_t(q_T) - g_t(q_0) - \int_0^T dt (D\nabla) \cdot \nabla g_t(q_t) \quad (4.17)$$

Stratonovitch integral Another stochastic integral that is useful is the Stratonovitch integral, denoted and defined by

$$\int_0^T dq_t^k \circ g_t(q_t) = \lim_{n\to\infty} \sum_{i=0}^{n-1} (q_{t_{i+1}}^k - q_{t_i}^k) g_t(\frac{q_{t_i} + q_{t_{i+1}}}{2}) \quad (4.18)$$

A nice property of the Stratonovitch integral is that it is antisymmetric with respect to time reversal, if q only contains positions, not velocities ($q = x$). When velocities are involved one also has to reverse the signs of the velocities, and the Stratonovitch integral may not be time-antisymmetric in this case. One can easily prove that for any function g

$$\int_0^T dq_t^k \circ g_t(q_t) = \int_0^T dq_t^k g_t(q_t) + \int_0^T dt (D\nabla)^k g_t(q_t) \quad (4.19)$$

so that the Stratonovitch integral actually has the normal integration rules:

$$\int_0^T [dq_t \circ \nabla g_t(q_t) + dt \frac{\partial g_t}{\partial t}(q_t)] = g_t(q_T) - g_t(q_0) \quad (4.20)$$

Work From mechanics we know that work should be equal to the integral over a path of force times displacement. But in diffusion systems this definition does not completely specify work, as one has to choose which stochastic integral to use. However, we also know that the work should be antisymmetric under time-reversal. Because of this, it can be seen that the work performed by a force $f_t(x)$ during ω is defined with a Stratonovitch integral:

$$\int_0^T dx_t \circ f_t(x_t) = \sum_k \int_0^T dx_t^k \circ f_t^k(x_t) \tag{4.21}$$

where x_t denotes a position, not a velocity.

4.7 Local detailed balance

Let us see what restrictions the local detailed balance assumption (Section 2.7) puts on our models.

Overdamped diffusions Let us consider general overdamped diffusions (4.13). One can then compute the quantity $S(\omega)$ (as defined in (2.12)) in the following way. The reference process, which is defined through (4.13) but with F put to zero, is a pure diffusion, which is an equilibrium process with the uniform distribution as equilibrium distribution. Actually in unbounded state spaces this distribution is not normalizable, so technically we do not have an equilibrium process. In any case we still have the important property that

$$\frac{d\mathcal{P}^0_{x_0}}{d\mathcal{P}^0_{x_T}\theta}(\omega) = 1$$

For the original process we can thus write

$$S(\omega) = \log \frac{d\mathcal{P}_{x_0}}{d\mathcal{P}^0_{x_0}}(\omega) - \log \frac{d\mathcal{P}_{x_T}\theta}{d\mathcal{P}^0_{x_T}\theta}(\omega) = A(\theta\omega) - A(\omega)$$

which after a short computation, using (4.15) and the definition of the Stratonovitch integral, becomes

$$S(\omega) = \int_0^T dx_t \circ D^{-1}(x_t)\chi(x_t)F_t(x_t)$$

By the local detailed balance assumption this should be equal to the entropy flux into the environment. Let us consider the case where the environment consists of

one heat bath at inverse temperature β. As an addition we write the force F_t as the sum of a conservative part $-\nabla U$ with U representing the energy of the system, and a nonconservative part f_t. Then we see that local detailed balance dictates that $\chi = \beta D$, because only in this case we have

$$S(\omega) = -\beta[U(x_T) - U(x_0) - \int_0^T dx_t \circ f_t(x_t)] \tag{4.22}$$

where in the brackets we have change of energy minus work done on the system (see (4.21)). This relation between mobility and the diffusion coefficient is called the Einstein relation. It is commonly used in Langevin dynamics, but here we see it as a consequence of the local detailed balance assumption.

Underdamped diffusions Let us return to the one-dimensional underdamped case (4.2). To compute $S(\omega)$ we should take into account that the velocities change sign under time-reversal:

$$S(\omega) = \frac{m\gamma}{\mathcal{D}} \int_0^T dt v_t F_t(x_t)$$

We can apply the same argument as for overdamped diffusions, and we see that local detailed balance dictates that $\mathcal{D} = \frac{m\gamma}{\beta}$, which is again called the Einstein relation.

Part II

Fluctuations and Response

"Recent developments in non equilibrium statistical physics have convinced us that times are ripe for a review of the vast subject concerning the fluctuations of systems described by statistical mechanics. This issue is important even beyond the 'traditional' applications of statistical mechanics, e.g. in a wide range of disciplines ranging from the study of small biological systems to turbulence, from climate studies to granular media etc. Moreover, the improved resolution in real experiments and the computational capability reached in numerical simulations has led to an increased ability to unveil the detailed nature of fluctuations, posing new questions and challenges to the theorists."

Umberto Marini Bettolo Marconi, Andrea Puglisi, Lamberto Rondoni and Angelo Vulpiani, in their introduction to *Fluctuation-Dissipation: Response Theory in Statistical Physics* (2008), [76]

Chapter 5

Response

In this part of the thesis, we investigate how a system responds to a perturbation, namely a small change in its energy. The central object that summarizes this is the response function. The goal is to relate this response function to correlation functions in the unperturbed process. In equilibrium systems this has already been done and the resulting relation is called the fluctuation-dissipation theorem. Out of equilibrium one needs to modify this relation. Extra terms appear that have to do with traffic, as defined in previous chapters. In this chapter we describe the general framework; in the next chapters more explicit models are examined. The research in this chapter was summarized earlier in [3].

5.1 Introduction

The fluctuation-dissipation theorem is a standard chapter in statistical mechanics [59, 76, 95]. In Chapter 2.10.4 we already explained that a small change in an equilibrium system that changes the Hamiltonian $H_0 \to H = H_0 - hV$ from time zero on, changes expectation values of observables $Q(x_t)$ at any time t. This change, up to linear order in h, is called the 'response' and is given by

$$\frac{\partial}{\partial h} \langle Q(x_t) \rangle_{\rho^0}^h \Big|_{h=0} = \beta \langle Q(x_t)[V(x_t) - V(x_0)] \rangle_\rho^0 \tag{5.1}$$

where the superscript h means that the average is taken in the perturbed system, and the superscript 0 stands for the unperturbed system, which is here an equilibrium system. This equality (5.1) is called the fluctuation-dissipation theorem. Actually it is usually stated in a different but equivalent form. In this

different form the parameter h is time-dependent: $h = h_t$, where $h_t = 0$ for $t < 0$. The time-dependence of h_t is something that is provided 'externally,' i.e. it is deterministic, it is controlled by the experimentalist. As an example, think of an Ising spin system, where the perturbation is a time-dependent magnetic field. The response is now the (functional) derivative of the expectation value of $Q(x_t)$ with respect to h_s with $s < t$. The fluctuation-dissipation theorem for equilibrium systems then states that

$$\frac{\delta}{\delta h_s} \langle Q(x_t) \rangle^h_{\rho^0} \Big|_{h=0} = \beta \frac{\partial}{\partial s} \langle Q(x_t) V(x_s) \rangle^0_{\rho^0} \tag{5.2}$$

We will derive this form of the fluctuation-dissipation theorem as a special case of the more general formula later on.

An early example of this theorem is present in Einstein's treatment of Brownian motion, where the diffusion constant, expressed as a velocity auto-correlation function, is found to be proportional to the mobility [33]. We will come back to this in the following chapters. Other famous examples include the Johnson-Nyquist formula for electronic white noise [55, 84] and the Onsager reciprocity for linear response coefficients [85, 86]. The fluctuation-dissipation theorem is useful because it gives a relation between two quantities in essentially different processes. One can e.g. determine the response of a system without actually perturbing it.

In this part of the thesis, we discuss and explain the results obtained in [2, 3, 4]. This research investigates linear response in systems out of equilibrium. So far, approaches deriving a fluctuation-dissipation relation (FDR) for systems out of equilibrium have not found a physical unification and do not appear as textbook material. One reason may be that previous work has not been seen to identify a sufficiently general structure with a clear corresponding statistical thermodynamic interpretation.

Aiming to provide a simple and general approach to FDR's, in [3] we have put forward a FDR for nonequilibrium regimes in a framework that may represent a unifying scheme for previous formulations. This is discussed in the remainder of this chapter. The application of this theory to Markov jump processes and overdamped diffusions was investigated in [4], and the application to underdamped diffusions in [2], and is explained in the next chapters. Additional work on fluctuation-dissipation relations for Markov jump processes was also reported in [75], but we do not discuss this here.

5.2 Linear Response

The unperturbed systems we consider are driven from equilibrium, but with a time-independent (time-homogeneous) dynamics. This dynamics can be far from equilibrium. We denote the energy of the system in a configuration x by $U(x)$. As a perturbation the energy is changed by the addition of a potential: $U \to U - h_t V$, where the time-dependent function h_t is the amplitude of the perturbation. We assume that it is bounded: $|h_t| < h < \infty$, with h a small number, and $h_t = 0$ for $t < 0$. We also assume that it is a continuous differentiable function of time. Furthermore, if the configuration of the system includes velocities, we always assume that energies and potentials that we use are symmetric under the sign-reversal of those velocities. With the notation of Section (2.5): e.g. $V(x) = V(\pi x)$.

We want to compute the influence of this perturbation on expectation values of observables $Q(x)$. We restrict ourselves to a regime where a small perturbation only has small consequences, which excludes for example phase-transitions. This regime allows us to make an expansion of expectation values in orders of the parameter h, which we do up to the first, linear, order:

$$\langle Q(x_t)\rangle^h_{\mu_0} = \langle Q(x_t)\rangle^0_{\mu_0} + \int_0^t ds h_s R_{QV}(t,s) + o(h) \tag{5.3}$$

which defines the response function R_{QV}, which actually also depends on the initial distribution μ_0:

$$R_{QV}(t,s) = \frac{\delta}{\delta h_s} \langle Q(x_t)\rangle^h_{\mu_0}\Big|_{h=0} \tag{5.4}$$

Applying the framework explained in Chapter 2: we use the action to rewrite expectation values of observables in the perturbed dynamics into expectation values in the unperturbed dynamics. Henceforth we denote by $\mathcal{P}^h$ the path-probability density of the perturbed process and by $\mathcal{P}^0$ that of the unperturbed one:

$$\langle Q(x_t)\rangle^h_{\mu_0} = \left\langle \frac{\mathcal{P}^h_{\mu_0}}{\mathcal{P}^0_{\mu_0}}(\omega) Q(x_t) \right\rangle^0_{\mu_0} = \left\langle e^{-A(\omega)} Q(x_t) \right\rangle^0_{\mu_0}$$

The averages are over all paths in the interval $[0,t]$. Because $h_t < h$ is small, the action is also small, and we can expand the exponential up to linear order in h:

$$\langle Q(x_t)\rangle^h_{\mu_0} = \langle Q(x_t)\rangle^0_{\mu_0} - \langle A(\omega) Q(x_t)\rangle^0_{\mu_0} + o(h)$$

Up to now this is only mathematics. Physics comes in when we split the action in its time-antisymmetric and time-symmetric parts (see Section 2.9):

$$\begin{aligned} S_{ex}(\omega) &= A(\theta\omega) - A(\omega) \\ \mathcal{T}_{ex}(\omega) &= A(\theta\omega) + A(\omega) \end{aligned}$$

where $S_{ex}(\omega)$ can, by the local detailed balance assumption, be interpreted as the excess entropy flux into the environment during the process ω. Excess is meant as excess of the perturbed process with respect to the unperturbed process. $\mathcal{T}_{ex}(\omega)$ is called the (excess) traffic. With this in mind we get

$$\langle Q(x_t)\rangle^h_{\mu_0} = \langle Q(x_t)\rangle^0_{\mu_0} + \frac{1}{2}\langle S_{ex}(\omega)Q(x_t)\rangle^0_{\mu_0} - \frac{1}{2}\langle \mathcal{T}_{ex}(\omega)Q(x_t)\rangle^0_{\mu_0} + o(h) \quad (5.5)$$

We have thus written the change in the expectation value of Q as the sum of the correlation functions of Q with entropy and traffic. This is equivalent to writing that the response function (5.4) equals

$$R_{QV}(t,s) = \frac{1}{2}\frac{\delta}{\delta h_s}\langle S_{ex}(\omega)Q(x_t)\rangle^0_{\mu_0}\Big|_{h=0} - \frac{1}{2}\frac{\delta}{\delta h_s}\langle \mathcal{T}_{ex}(\omega)Q(x_t)\rangle^0_{\mu_0}\Big|_{h=0} \quad (5.6)$$

This constitutes our most general but also our most vague result: the response function is the sum of two terms. One of the terms is expressed using the entropy flux into the environment, which is in principle known and measurable. The other term is up to here only mathematically defined and needs more investigation. Before we do that, we first discuss some general properties of the response function.

5.3 General properties of the response function

5.3.1 Relaxation

We assume that the unperturbed dynamics is time-homogeneous. Then, as we assume in this thesis, the probability distribution of states relaxes to the stationary distribution, meaning that

$$\langle Q(x_t)\rangle^0_{\mu_0} \to \langle Q(x_t)\rangle^0_{\rho} \qquad \text{for } t \to \infty$$

Suppose now that the perturbation we added to the system, is removed again after a time t_1, i.e. $h_t = 0$ for $t > t_1$. Then logically the system relaxes again to the stationary regime:

$$\langle Q(x_t)\rangle^h_{\mu_0} \to \langle Q(x_t)\rangle^0_{\rho} \qquad \text{for } t \to \infty$$

Using (5.3), this gives

$$\int_0^{t_1} ds h_s R_{QV}(t,s) \to 0 \qquad \text{for } t \to \infty$$

As this is true for any h_t, we can safely say that $R_{QV}(t,s) \to 0$ for $t \to \infty$. However, throughout this text, we make a stronger assumption, namely that

$$\int_s^{\infty} dt |R(t,s)| < \infty \tag{5.7}$$

for any s. The reason for this assumption is that the excess heat dissipation remains finite, see Section 5.4.3.

5.3.2 Causality

Physically, the response function should be zero whenever $s > t$. This is a consequence of causality, because the expectation value of an observable at one time can not be influenced by a perturbation at a later time. What we want to show here, is that causality is already embedded in our framework, and not something we have to put in afterwards by hand. In the discussion after (2.4) we argued that to take the expectation value of $Q(x_t)$, we only need to average over paths in the interval $[0,t]$. The perturbed dynamics that govern the probability of this path, and thus also the action $A(\omega)$, can logically only depend on h_s with $0 \leq s \leq t$. So the perturbed expectation value of $Q(x_t)$ is independent of h_s with $s > t$. This is indeed causality. In (5.5) causality gives us that for $s > t$:

$$\frac{\delta}{\delta h_s} \langle S_{ex}(\omega) Q(x_t) \rangle_{\mu_0}^0 \Big|_{h=0} = \frac{\delta}{\delta h_s} \langle \mathcal{T}_{ex}(\omega) Q(x_t) \rangle_{\mu_0}^0 \Big|_{h=0} \tag{5.8}$$

5.3.3 A constant perturbation

For experiments and simulations, it is often more convenient to work with a constant perturbation. This means that $h_s = h$ for $s \geq 0$ and is equal to zero for $s < 0$. In this case one will not directly measure the response function, but rather the time-integral of it. This is because (5.3) reduces in this case to:

$$\langle Q(x_t) \rangle_{\mu_0}^h = \langle Q(x_t) \rangle_{\mu_0}^0 + h \int_0^t ds R_{QV}(t,s) + o(h)$$

and the response is then rather defined as

$$\frac{\partial}{\partial h}\left\langle Q(x_t)\right\rangle_{\mu_0}^{h}\Big|_{h=0} = \int_0^t ds R_{QV}(t,s)$$

This integrated response is sometimes called a generalized susceptibility, as a generalization of the susceptibility defined in thermodynamics.

5.3.4 Response in the frequency domain

Apart form a constant perturbation, a periodic perturbation is also often convenient in experimental settings, and response functions are often measured in the frequency domain (using Fourier transforms). Let us consider therefore a perturbation with an amplitude $h_t = h_0 \cos(kt)$. We also restrict ourselves to the case that the unperturbed system before the perturbation was stationary:

$$\left\langle Q(x_t)\right\rangle_{\rho}^{h} = \left\langle Q(x)\right\rangle_{\rho}^{0} + h_0 \int_0^t ds \cos(ks) R_{QV}(t-s) + o(h_0) \tag{5.9}$$

Up to now we have always taken $t = 0$ as the moment that the perturbation was turned on. If we change this time to $t = -T$, then it only changes the lower limit of the integral in (5.9) from 0 to $-T$. We suppose that the perturbation was taken a long time ago, so we can effectively take $-T \to -\infty$. Our assumption (5.7) ensures that this makes sense. The upper limit of the integral (5.9) can also be changed to $+\infty$, as $R_{QV}(t-s) = 0$ for $s > t$ due to causality. This naturally leads us to rewriting (5.9) in terms of Fourier transforms:

$$\begin{aligned}\int_{-\infty}^{+\infty} ds \cos(ks) R_{QV}(t-s) &= \int_{-\infty}^{+\infty} ds \frac{e^{iks}+e^{-iks}}{2} R_{QV}(t-s) \\ &= \frac{1}{2}\left[\tilde{R}_{QV}(k)e^{-ikt} + \tilde{R}_{QV}(-k)e^{ikt}\right]\end{aligned}$$

where the Fourier transform of the response function is defined as $\tilde{R}_{QV}(k) = \int_{-\infty}^{+\infty} dt R_{QV}(t) e^{ikt}$. Using properties of Fourier transforms, and the fact that $R_{QV}(t)$ is a real-valued function, we arrive at

$$\left\langle Q(x_t)\right\rangle_{\rho}^{h} = \left\langle Q(x)\right\rangle_{\rho}^{0} + h_0 \mathrm{Re}(\tilde{R}_{QV}(k)e^{-ikt}) + o(h_0)$$

where $\mathrm{Re}(z)$ is the real part of a complex number z. So for these kinds of perturbations, what is actually measured is the Fourier transform of the response function.

Using complex analysis, one can prove [78] that $\tilde{R}_{QV}(k)$ is a continuous bounded function, as a consequence of the assumption (5.7).

5.4 The case of one heat bath

The result (5.5) is a very general, but also a very vague result. To get more physical results we can of course investigate explicit examples, which we do in the next chapters. Before doing this however, we write down a more explicit form for the response function by utilizing the fact that entropy is a known physical quantity. To do this we restrict ourselves to systems in contact with an environment at a single temperature. This means the system can be driven away from equilibrium by a nonconservative force or by particle reservoirs at different chemical potentials. Throughout the rest of this chapter and the next, we keep this assumption. In the example of underdamped diffusions we will come back to the case of heat baths at different temperatures.

Entropy, work and heat: In the case that the environment of the system is described by a single temperature, we can write the entropy flux into the environment as

$$S(\omega) = \beta Q(\omega) = \beta[-\Delta E + W(\omega)]$$

with Q the heat flow into the environment, ΔE the change of energy of the system, and W the work done on the system. The excess entropy flux is the extra entropy flux created by the addition of the potential $-h_t V$, and is thus equal to

$$S_{ex}(\omega) = \beta Q_{ex}(\omega) = \beta[h_t V(x_t) - h_0 V(x_0) - \int_0^t ds \frac{\partial h_s}{\partial s} V(x_s)] \tag{5.10}$$

where the last term is the work due to the added potential $-h_t V$. We give a short argument to see why this is work: divide the time interval $[0, t]$ into n segments of length Δt, and define $t_i = i\Delta t$ for $i = 1, \ldots, n$. We then have that

$$\begin{aligned} h_t V(x_t) - h_0 V(x_0) &= \sum_{i=1}^{n} [h_{t_i} V(x_{t_i}) - h_{t_{i-1}} V(x_{t_i})] \\ &\quad + \sum_{i=1}^{n} [h_{t_{i-1}} V(x_{t_i}) - h_{t_{i-1}} V(x_{t_{i-1}})] \end{aligned} \tag{5.11}$$

The left-hand side of this equation is a change of energy, and can thus be split into work and heat. In our discussion in Section 1.2 we defined heat as the work due

to (stochastic) forces we do not detect, and can not be externally controlled. If we apply this definition here, we can see that the second term on the right-hand side exists due to changes in the configuration, which are stochastic and can't be controlled. The first term is due to changes in the parameter h_t, which is controlled externally. Therefore the first term in (5.11) is interpreted as work and the second as heat (see also [19]). Taking the limit of $n \to \infty$ for the work-term, we get exactly the last term in (5.10). Note that we can't write the heat in terms of a partial derivative of V with respect to x, because that derivative usually does not exist, or even makes sense (e.g. for Markov jump processes).

Response function With this knowledge we can rewrite the correlation of the entropy flux with the observable as follows:

$$\begin{aligned}\langle S_{ex}(\omega)Q(x_t)\rangle^0_{\mu_0} &= \beta[h_t \langle V(x_t)Q(x_t)\rangle^0_{\mu_0} - h_0 \langle V(x_0)Q(x_t)\rangle^0_{\mu_0} \\ &\quad - \int_0^t ds \frac{\partial h_s}{\partial s} \langle V(x_s)Q(x_t)\rangle^0_{\mu_0}] \\ &= \beta \int_0^t ds h_s \frac{\partial}{\partial s} \langle V(x_s)Q(x_t)\rangle^0_{\mu_0} \end{aligned} \tag{5.12}$$

To write down the response function (5.4), it is convenient to define the functional derivative of the traffic:

$$\tau(\omega, s) = \frac{\partial}{\partial h_s} \mathcal{T}_{ex}(\omega)\Big|_{h=0} \tag{5.13}$$

With this notation and using (5.12) the response function (5.4) becomes:

$$R_{QV}(t,s) = \frac{\beta}{2}\frac{\partial}{\partial s} \langle V(x_s)Q(x_t)\rangle^0_{\mu_0} - \frac{1}{2} \langle \tau(\omega, s)Q(x_t)\rangle^0_{\mu_0} \tag{5.14}$$

This is the most general formula for Markovian systems out of equilibrium, but in contact with only one heat bath at inverse temperature β. We see that the first term on the right-hand side coincides formally with one half of the response function for equilibrium systems. It is of course the second term that is different for nonequilibrium systems with respect to equilibrium. For a constant perturbation $h_t = h$, we get

$$\int_0^t ds R_{QV}(t,s) = \frac{\beta}{2} \langle [V(x_t) - V(x_0)]Q(x_t)\rangle^0_{\mu_0} - \frac{1}{2}\int_0^T ds \langle \tau(\omega, s)Q(x_t)\rangle^0_{\mu_0}$$

5.4.1 Causality and Stationarity

For the case of one heat bath, we can combine (5.14) with (5.8) to find for $t > s$:

$$\beta \frac{\partial}{\partial t} \langle V(x_t) Q(x_s) \rangle^0_{\mu_0} = \langle \tau(\omega, t) Q(x_s) \rangle^0_{\mu_0} \tag{5.15}$$

We can use this causality relation to rewrite (5.14) for $t > s$:

$$\begin{aligned} R_{QV}(t,s) &= \frac{\beta}{2} \left[\frac{\partial}{\partial s} \langle V(x_s) Q(x_t) \rangle^0_{\mu_0} - \frac{\partial}{\partial t} \langle V(x_t) Q(x_s) \rangle^0_{\mu_0} \right] \\ &\quad + \frac{1}{2} \left[\langle \tau(\omega, t) Q(x_s) \rangle^0_{\mu_0} - \langle \tau(\omega, s) Q(x_t) \rangle^0_{\mu_0} \right] \end{aligned} \tag{5.16}$$

Consider now the case that the system was in the stationary state when the perturbation was added, meaning $\mu_0 = \rho$ is the stationary distribution. Stationarity of the process means that it is time-translation invariant. As a consequence, the response function can be written as $R_{QV}(t,s) = R_{QV}(t-s)$. Moreover, correlation functions computed in the stationary regime satisfy

$$\langle V(x_s) Q(x_t) \rangle_\rho = \langle V(x_0) Q(x_{t-s}) \rangle_\rho$$

This means that in this case (5.16) simplifies to

$$\begin{aligned} R_{QV}(t,s) &= \frac{\beta}{2} \frac{\partial}{\partial s} \left[\langle V(x_s) Q(x_t) \rangle^0_\rho + \langle V(x_t) Q(x_s) \rangle^0_\rho \right] \\ &\quad + \frac{1}{2} \left[\langle \tau(\omega, t) Q(x_s) \rangle^0_\rho - \langle \tau(\omega, s) Q(x_t) \rangle^0_\rho \right] \end{aligned} \tag{5.17}$$

valid for $t > s$.

5.4.2 Equilibrium

We prove here the fluctuation-dissipation relation for equilibrium systems. A system in equilibrium is necessarily in contact with only one heat bath, so (5.14) applies here. A system in equilibrium is also time-reversible. The time-reversal operator is defined as $\theta\omega = (\pi x_{t-s})_{0 \le s \le t}$, where π reverses the sign of velocities. Time-reversibility means that the expectation values of observables are equal to the expectation values of their time-reversed twins, e.g:

$$\langle V(x_s) Q(x_t) \rangle^0_\rho = \langle V(x_t) Q(\pi x_s) \rangle^0_\rho$$

where we used the assumption that $V(x) = V(\pi x)$. From the definition of $\tau(\omega, s)$ (see (5.13)), it also follows that $\tau(\theta\omega, s) = \tau(\omega, t-s)$, so that for $t > s$:

$$\langle \tau(\omega, s) Q(x_t) \rangle_\rho^0 = \langle \tau(\omega, t-s) Q(\pi x_0) \rangle_\rho^0 = \langle \tau(\omega, t) Q(\pi x_s) \rangle_\rho^0$$

which by the causality relation (5.15) equals

$$\langle \tau(\omega, t) Q(\pi x_s) \rangle_\rho^0 = \beta \frac{\partial}{\partial t} \langle V(x_t) Q(\pi x_s) \rangle_\rho^0 = -\beta \frac{\partial}{\partial s} \langle V(x_0) Q(x_{t-s}) \rangle_\rho^0$$

Substituting this equality in (5.16) we arrive at

$$R_{QV}(t, s) = \beta \frac{\partial}{\partial s} \langle V(x_s) Q(x_t) \rangle_\rho^0 \qquad \text{for} \quad t > s$$

which is exactly the fluctuation-dissipation theorem in equilibrium. Note that for a constant perturbation $h_t = h$ the response becomes

$$\int_0^t ds R_{QV}(t, s) = \beta \langle [V(x_t) - V(x_0)] Q(x_t) \rangle_\rho^0$$

as in (5.1).

5.4.3 Dissipation

In equilibrium it is known that the response function is closely related to the energy dissipation of the system into the environment [78]. This is already seen in the fact that the response function is then expressed through the fluctuation-dissipation theorem as a correlation functional between the observable and entropy flux. In systems driven out of equilibrium things get more complicated: even when the system is not perturbed there is already a heat dissipation. This is called the 'housekeeping' heat that is needed to maintain the (unperturbed) nonequilibrium stationary state. Perturbing the system then gives additional heat. We see here that the prediction in [88] that the usual equilibrium relation between response and dissipation is preserved when taking into account only the excess heating and ignoring the housekeeping heat, is indeed true in the following sense.

For a system in contact with only one heat bath, the entropy flux into the environment is equal to β times the heat dissipated into the environment. For each trajectory ω, this heat can be split in the heat of the unperturbed process and the excess heat of the perturbed process with respect to the unperturbed

process. In (5.10) we already wrote down explicitly the excess heat. We assume a perturbation of the form $h_t = h_0 \cos(kt)$, like we did in Section (5.3.4), and compute the excess heat over one period of the perturbation $[0, T = \frac{2\pi}{k}]$ in the perturbed system:

$$\begin{aligned}\langle Q_{ex}(\omega)\rangle_\rho^h &= h_T \langle V(x_T)\rangle_\rho^h - h_0 \langle V(x_0)\rangle_\rho^h - \int_0^T dt \frac{\partial h_t}{\partial t} \langle V(x_t)\rangle_\rho^h \\ &= \int_0^T dt h_t \frac{\partial}{\partial t} \langle V(x_t)\rangle_\rho^h\end{aligned}$$

We can write the expectation of V using the response function as in (5.10):

$$\langle V(x_t)\rangle_\rho^h = \langle V(x)\rangle_\rho^0 + \frac{h_0}{2}\left[\tilde{R}_{VV}(k)e^{-ikt} + \tilde{R}_{VV}(-k)e^{ikt}\right]$$

Substituting this, and the explicit expression of h_t into the expression of the heat, gives us

$$\begin{aligned}\langle Q_{ex}(\omega)\rangle_\rho^h &= \frac{ikh_0^2}{2}\int_0^T dt \cos(kt)\left[-\tilde{R}_{VV}(k)e^{-ikt} + \tilde{R}_{VV}(-k)e^{ikt}\right] + o(h_0^2) \\ &= \frac{ikh_0^2 T}{4}\left[\tilde{R}_{VV}(-k) - \tilde{R}_{VV}(k)\right] + o(h_0^2) \\ &= \pi h_0^2 \mathrm{Im}[\tilde{R}_{VV}(k)] + o(h_0^2)\end{aligned}$$

We see that the excess dissipated heat can be expressed through the imaginary part of the Fourier transform of the response function. For systems that are (unperturbed) in equilibrium, there is no housekeeping heat, so the excess heat is then total dissipated heat. Although the excess dissipated heat only specifies the imaginary part, the Kramers-Kronig relations in complex analysis connect the real and imaginary parts of complex functions to each other [78]. This means that when one of them is known, the other can be computed from it.

Let us conclude here by noting that the imaginary part of the Fourier transform is equal to the Fourier transform of the time-antisymmetric part of the response function (5.14):

$$\begin{aligned}r(t-s) &= \frac{1}{2}[R_{VV}(t,s) - R_{VV}(s,t)] \\ &= \beta \frac{\partial}{\partial s}\langle V(x_s)V(x_t)\rangle_\rho^0 + \frac{1}{2}\left[\langle \tau(\omega,t)V(x_s)\rangle_\rho^0 - \langle \tau(\omega,s)V(x_t)\rangle_\rho^0\right]\end{aligned}$$

which is the same as (5.17) for $Q = V$, but (5.17) is only valid for $t > s$. For $t > s$, we have that $r(t-s) = R_{VV}(t-s)$, while for $t < s$ the response function is zero. The antisymmetric part of the response function therefore contains all information of the whole response function.

Chapter 6

Markov jump processes and overdamped diffusions

In this chapter the general framework outlined in the previous chapter is applied to the cases of Markov jump processes and overdamped diffusions. These two classes of models share a common property, simplifying the fluctuation-dissipation relation, as we shall see in (6.7). For a few explicit examples, simulations have been done to visualize the terms in the fluctuation-dissipation relations. The work explained here was written down in [4].

6.1 An explicit formula

In this chapter we only consider systems in contact with one heat bath, so that we can use formulae like (5.14) and (5.16). The advantage of considering specific models is that we can get more explicit results. To do this, we need to compute the traffic, i.e. the time-symmetric part of the action for our models. More precisely, we need to compute the functional derivative of traffic, as defined in (5.13)

6.1.1 Overdamped diffusions

Consider the case of general overdamped diffusions in d dimensions (4.13). Adding the potential $-h_t V(x_t)$ in Langevin equations is straightforward. The perturbed Langevin equation becomes (we are still working in the Itô convention):

$$dx_t = \chi(x_t)[F(x_t) + h_t \nabla V(x_t)]dt + \nabla \cdot D(x_t)dt + \sqrt{2D(x_t)}dB_t$$

Remember that the local detailed balance assumption here implies that $\chi = \beta D$. In the same way as we derived (4.15), we can here derive the action that describes the relative probability of paths of the perturbed process with respect to the unperturbed process:

$$\begin{aligned} A(\omega) &= -\log \frac{d\mathcal{P}^h}{d\mathcal{P}^0}(\omega) \\ &= -\frac{\beta}{2}\int_0^t dx_s h_s \nabla V + \frac{\beta}{2}\int_0^t ds h_t \nabla V \cdot [\nabla \cdot D + \chi F + \frac{1}{2} h_t \chi \nabla V] \end{aligned}$$

The time-antisymmetric part of the action is given by a Stratonovitch integral:

$$S_{ex}(\omega) = A(\theta\omega) - A(\omega) = \beta \int_0^t dx_s \circ h_s \nabla V$$

as a consistency check, let us rewrite this using (4.20):

$$\int_0^t dx_s \circ h_s \nabla V = h_t V(x_t) - h_0 V(x_0) - \int_0^t ds \frac{\partial h_s}{\partial s} V(x_s)$$

which is exactly (5.10). More importantly, we can compute explicitly the time-symmetric part of the action:

$$\mathcal{T}_{ex}(\omega) = A(\theta\omega) + A(\omega) = \beta \int_0^t ds h_s [\nabla \cdot (D\nabla) + \chi F \nabla + \frac{1}{2} h_s (\nabla V) \chi \nabla] V$$

so that the functional derivative with respect to h_s becomes:

$$\tau(\omega, s) = \beta \frac{\partial}{\partial h_s} \mathcal{T}_{ex}(\omega)\Big|_{h=0} = [\nabla \cdot (D(x_s)\nabla) + \chi(x_s) F(x_s) \nabla] V(x_s) \qquad (6.1)$$

and the fluctuation-dissipation relation:

$$\begin{aligned} R_{QV}(t,s) &= \frac{\beta}{2}\frac{\partial}{\partial s} \langle V(x_s) Q(x_t) \rangle_{\mu_0}^0 \qquad (6.2) \\ &\quad -\frac{\beta}{2} \langle [\nabla \cdot (D(x_s)\nabla V(x_s)) + \chi(x_s) F(x_s) \nabla V(x_s)] Q(x_t) \rangle_{\mu_0}^0 \end{aligned}$$

6.1.2 Markov jump processes

For Markov jump processes, it is more ambiguous to define what 'adding a potential' means. Starting from a general Markov jump process with transition rates $k(x,y)$, the local detailed balance assumption asserts that (see (3.10)):

$$\frac{k(x,y)}{k(y,x)} = e^{-\beta[U(y)-U(x)-W(x,y)]}$$

where U is the energy of the system, and $W(x,y)$ is the work needed for a jump $x \to y$. Adding a potential $-h_t V$ changes the transition rates to new time-dependent rates $k_t(x,y)$ which, again by the local detailed balance assumption, have to satisfy

$$\frac{k_t(x,y)}{k_t(y,x)} = e^{-\beta[U(y)-U(x)-W(x,y)-h_t V(y)+h_t V(x)]}$$

Still, this does not completely specify the perturbation. There are many possibilities. The most common in literature however is the following one, which we use throughout this chapter:

$$k_t(x,y) = k(x,y) e^{\frac{\beta h_t}{2}[V(y)-V(x)]}$$

For the treatment of fluctuation-dissipation relations for more general perturbations, see [26, 75, 92]. From (3.7) we arrive straightforwardly at

$$\begin{aligned} A(\omega) &= -\log \frac{d\mathcal{P}^h}{d\mathcal{P}^0}(\omega) \\ &= -\frac{\beta}{2}\sum_{s\le t} h_s[V(x_s) - V(x_{s^-})] \\ &\quad + \int_0^t ds \sum_y k(x_s,y)[e^{\frac{\beta h_t}{2}[V(x_s)-V(x_{s^-})]} - 1] \end{aligned}$$

The time-antisymmetric part of the action is here

$$S_{ex}(\omega) = \beta \sum_{s\le t} h_s[V(x_s) - V(x_{s^-})]$$

one can check that this is consistent with (5.10). The traffic is the time-symmetric part of the action:

$$\mathcal{T}_{ex}(\omega) = A(\theta\omega) + A(\omega) = 2\int_0^t ds \sum_y k(x_s, y)[e^{\frac{\beta h_t}{2}[V(y)-V(x_s)]} - 1]$$

and its functional derivative:

$$\tau(\omega, s) = \frac{\partial}{\partial h_s}\mathcal{T}_{ex}(\omega)\Big|_{h=0} = \beta \sum_y k(x_s, y)[V(y) - V(x_s)] \tag{6.3}$$

The fluctuation-dissipation relation then becomes

$$\begin{aligned} R_{QV}(t,s) &= \frac{\beta}{2}\frac{\partial}{\partial s}\langle V(x_s)Q(x_t)\rangle_{\mu_0}^0 \\ &\quad -\frac{\beta}{2}\left\langle \sum_y k(x_s, y)[V(y) - V(x_s)]Q(x_t)\right\rangle_{\mu_0}^0 \end{aligned} \tag{6.4}$$

6.1.3 A general formula

An important property of both Markov jump processes and overdamped diffusions, is that the functional derivative of traffic is a state function, see (6.1) and (6.3), i.e. $\tau(\omega, s) = \tau(x_s)$. With this in mind we use the equality (5.15):

$$\beta\frac{\partial}{\partial s}\langle V(x_s)Q(x_t)\rangle_{\mu_0}^0 = \langle \tau(x_s)Q(x_t)\rangle_{\mu_0}^0 \tag{6.5}$$

which is true for any $s > t$, Q and μ_0. We use the (backward) generator of the unperturbed Markov process, which is defined (see Appendix A) as follows: there exists an operator L such that for any state function f and any initial distribution μ_0:

$$\frac{d}{dt}\langle f(x_t)\rangle_{\mu_0}^0 = \langle Lf(x_t)\rangle_{\mu_0}^0$$

This L is called the backward generator of the (unperturbed) Markov process. If we use this generator and take $Q = 1$ in (6.5) then we get:

$$\langle \tau(x_s)\rangle_{\mu_0}^0 = \beta\langle LV(x_s)\rangle_{\mu_0}^0$$

and because this is true for any s and any μ_0, we actually have that

$$\tau(x_s) = \beta LV(x_s) \tag{6.6}$$

Comparing (6.1) and (6.3) to (A.8) and (A.11) we can check that this is indeed the case. This expression also gives an interpretation of $\tau(x)$. Using the definition of the generator L, we see that

$$\tau(x) = \beta LV(x) = \beta \frac{d}{dt} \langle V(x_t) \rangle_x^0 \Big|_{t=0}$$

where the superscript x means that the initial configuration is fixed at x. Thus $\tau(x)$ is the instantaneous expected change in the potential V when started at the state x.

The fluctuation-dissipation relation becomes :

$$R_{QV}(t,s) = \frac{\beta}{2} \frac{\partial}{\partial s} \langle V(x_s) Q(x_t) \rangle_{\mu_0}^0 - \frac{\beta}{2} \langle LV(x_s) Q(x_t) \rangle_{\mu_0}^0 \tag{6.7}$$

which is the most general formulation when restricted to Markov jump processes and overdamped diffusions. These are large and useful classes of models. Moreover, notice that the result is true for any observable $Q(x_t)$, any potential V and any initial distribution μ_0. The result is thus true in three nonequilibrium situations:

1. The case in which there are nonconservative forces.
2. The case in which there are different particle reservoirs at the same temperature but with different chemical potentials.
3. The case in which the system has an equilibrium dynamics, but has not yet relaxed to equilibrium.

Of course combinations of these nonequilibrium conditions are also possible. In this sense the result (6.7) is both very general, and explicit in terms of the generator of the process. The downside is that experimentally, this generator is not always known. On the other hand, in the next section we show some examples where this formula leads to (in principle) measurable results.

6.2 Examples

In this section we give three physical examples to clarify the structure of the fluctuation-dissipation relations. For two of these examples, simulations have been

made by Marco Baiesi, to provide some visualization and a check of the results [4]. For simulations it is often more convenient to use a constant perturbation: $h_s = h, s \geq 0$. As discussed in Section 5.3.3, this gives an integrated version of the fluctuation-dissipation relation. For this section, it is convenient to introduce the following notation:

- The generalized susceptibility:

$$\chi(t) = \lim_{h\to 0} \frac{1}{h} \left[\langle Q(x_t) \rangle^h_{\mu_0} - \langle Q(x_t) \rangle^0_{\mu_0} \right]$$

  If extended to $t \uparrow \infty$, $\chi(t)$ gives the change in nonequilibrium stationary expectation when adding a small potential.

- The correlation function, originally coming from the entropic term in (6.7):

$$C(t) = \langle Q(x_t) V(x_t) \rangle^0_{\mu_0} - \langle Q(x_t) V(x_0) \rangle^0_{\mu_0}$$

- The term coming from the traffic term in (6.7), (extra with respect to equilibrium):

$$K(t) = -\int_0^t ds \, \langle LV(x_s) Q(x_t) \rangle^0_{\mu_0}$$

  representing an integrated correlation function.

- The average of C and K:

$$C_{NE}(t) = \frac{1}{2}[C(t) + K(t)]$$

The fluctuation-dissipation relation is in this notation expressed by

$$\chi(t) = \beta C_{NE}(t) \tag{6.8}$$

6.2.1 Driven Kawasaki dynamics

As a first example we consider an exclusion process as a model of ionic transport through a narrow channel. The model we use is very similar to the exclusion process described as example 2 in Section 3.8, differing only by the addition of an interaction potential. We repeat quickly: the model is described by a collection of n sites, labelled by $i = 1, \ldots, n$, each holding either one particle ($x^i = 1$) or none ($x^i = 0$). A configuration x is thus an array of ones and zeros. In the bulk of this system no particles are created or annihilated, only jumping to neighbouring sites

is allowed (Kawasaki dynamics). At the edges $i = 1, n$ particles can move in or out from reservoirs with density d_1 and d_n, respectively. To this a nearest neighbour interaction is added, with energy $U(x) = -\sum_{i=1}^{n} x^i x^{i+1}$. Moreover, we can add an "electric" field E promoting particle jumps to the right.

We have to construct transition rates for particles hopping to neighbouring sites and rates for creation and annihilation at the edges. For example, consider the case that a particle enters into site $i = 1$ from the reservoir. Denote the configuration before that jump x, and after the jump y, where $y = x$ except that $x^1 = 0$ while $y^1 = 1$. Using the local detailed balance assumption, we know that

$$\frac{k(x,y)}{k(y,x)} = e^{\beta\mu_1} e^{-\frac{\beta}{2}[U(y)-U(x)]} = \frac{d_1}{1-d_1} e^{-\frac{\beta}{2}[U(y)-U(x)]}$$

This does not specify the rates fully of course. For simulational purposes we therefore define the rate for a particle entering at $i = 1$ as

$$k(x,y) = d_1 \psi(x,y) \exp\left\{-\frac{\beta}{2}[U(y) - U(x)]\right\}, \quad \psi(x,y) = \psi(y,x)$$

where for the moment $\psi(x,y)$ is an arbitrary function. Then of course, a particle leaving gives

$$k(y,x) = (1-d_1)\psi(x,y) \exp\left\{\frac{\beta}{2}[U(y) - U(x)]\right\}$$

and similar for particles entering and leaving at site n. In the bulk of the system particles can hop, e.g. to the right. Let x and y denote configurations which differ by the fact that a particle at site i has hopped to the right: $x^i = 1, x^{i+1} = 0$ and $y^i = 0, y^{i+1} = 1$. Then local detailed balance gives

$$\frac{k(x,y)}{k(x,y)} = \exp\{-\beta[U(y) - U(x) - E]\}$$

Where E is an electric field promoting hops to the right. We therefore take the hopping rates to be:

$$k(x,y) = \psi(x,y) \exp\left\{-\frac{\beta}{2}[U(y) - U(x) - E]\right\}, \quad \psi(x,y) = \psi(y,x)$$

and similarly for jumps to the left. For simulational purposes we have specified $\psi(x,y)$ by an additional condition, namely that $k(x,y) + k(y,x) = 1$ for all x, y.

The system is driven from equilibrium by:

i) setting different reservoir densities $d_1 \neq d_n$, or

ii) setting a nonzero electric field $E > 0$ in the bulk.

We choose the total number of particles $\mathcal{N}(t) = \sum_{i=1}^{n} x^i$ as observable. We also introduce a perturbation $V(s)$ equal to $\mathcal{N}(s)$. This particular perturbation is equivalent to changing the chemical potential of both reservoirs with a common shift: $\mu_i \to \mu_i + h$:

$$k_h(x,y) = k(x,y)e^{-\frac{\beta h}{2}[\mathcal{N}(y)-\mathcal{N}(x)]}$$

Transition rates for the perturbed process are thus multiplied by a factor $e^{\beta h/2}$ if a particle enters the system, and by $e^{-\beta h/2}$ when a particle leaves; transitions in the bulk are left unchanged.

In this case the traffic term is

$$L\mathcal{N}(x) = \sum_y k(x,y)[\mathcal{N}(y) - \mathcal{N}(x)] = \mathcal{J}(x)$$

which is a current. This current represents the expected change of $\mathcal{N}$ per unit time, i.e., it is the rate of change in the number of particles from the two possible transitions at the boundary sites. Thus, for $x \to y$ the transition modifying x^1, and for $x \to z$ the transition modifying x^n, we have

$$\mathcal{J}(x) = [\mathcal{N}(y) - \mathcal{N}(x)]k(x,y) + [\mathcal{N}(z) - \mathcal{N}(x)]k(x,z)$$

We have numerically verified that χ, C, K as defined earlier, are all equal to each other under equilibrium conditions. While $C = K$ to excellent precision, the shape of χ depends weakly on h, and is found to converge to C only for h sufficiently small. In fact, one can pretend exact matching only in the limit $h \to 0$, but $h = 0.01$ turns out to be sufficiently small to achieve a good convergence.

A representative example for the nonequilibrium case i) is shown in Fig. 6.1(a). One can see that each of the functions $C(t)$ and $K(t)$ is a poor approximation of the response, while the agreement between $C_{NE}(t)$ and $\chi(t)/\beta$ is excellent.

In Fig. 6.1(b) we show an example of nonequilibrium condition ii). Again, only $C_{NE}(t)$ matches $\chi(t)/\beta$. Curiously, a comparison of this example with the previous one reveals that C can be either larger or smaller than K, even for two nonequilibrium conditions that look pretty similar, in the sense that they yield a current in the same direction for a relatively simple system.

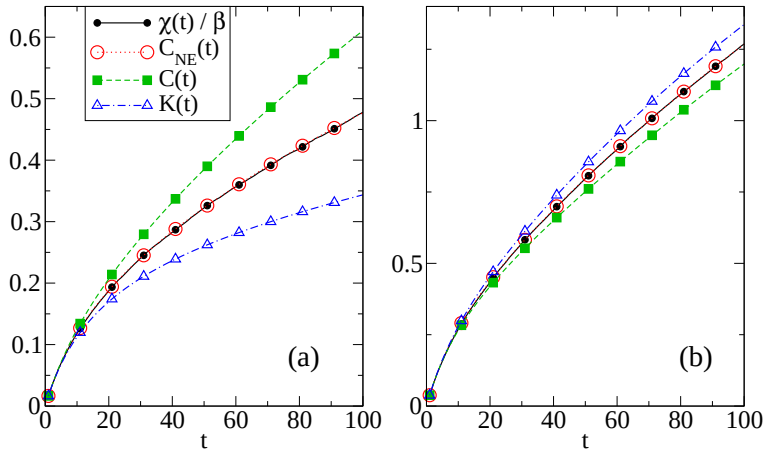


Figure 6.1: Plot of the quantities involved in Eq. (6.8), for (a) case i) ($E = 0$) with $n = 10$, $\beta = 1$, $h = -0.01$, and reservoir density unbalance $d_1 = 0.9$, $d_n = 0.1$, and (b) for case ii) ($d_1 = d_n = 0.5$) with $n = 10$, $\beta = 1$, $h = -0.01$, $E = 3$.

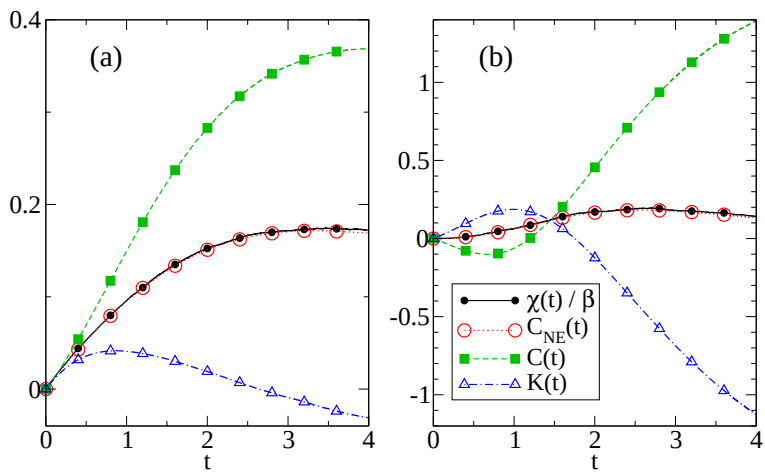


Figure 6.2: Response and fluctuations of the overdamped particle in a tilted periodic potential, as discussed in the text, with inverse temperature $\beta = 0.2$, mobility $\nu = 1$ and perturbation $h = 0.02$. (a) Steady state regime, with initial distribution equal to the stationary one, and force $f = 0.9$. (b) Transient regime, with initial position $x_0 = 0$ and force $f = 0$.

6.2.2 Diffusion on the circle

Consider the overdamped Langevin equation (4.10) for a particle position $x_t \in S^1$ (on a circle) with a force $F(x) = f - U'(x)$. Here the prime denotes differentiation with respect to space. The constant force f drives the particle around the circle and out of equilibrium. To this equation a perturbation is added:

$$dx_t = \nu[f - U'(x_t) + h_t V'(x_t)]dt + \sqrt{2D}dB_t \tag{6.9}$$

we renamed the mobility ν to avoid confusion with the integrated response $\chi(t)$. The diffusion constant D and the mobility ν are related by the Einstein relation $\nu = \beta D$, with β the inverse temperature (local detailed balance).

This example has been recently experimentally realized as reported in [45], for testing the results of [15, 16], see also [8, 97]. Moreover, its nonequilibrium stationary distribution ρ is known analytically, see [71].

Here we show the result of a simple simulation of the overdamped particle studied in the experiment, with a potential $V(x) = \sin(x)$, and also $U(x) = Q(x) = V(x)$, and a constant force f. The functional derivative of the traffic $\tau(x) = \beta LV(x)$ of a particle at position x can be computed by applying the generator of the overdamped dynamics

$$L = (f - U')\frac{d}{dx} + \frac{1}{\beta}\frac{d^2}{dx^2}$$

to the potential $V(x)$, which gives $\tau(x) = \beta[f - \cos(x)]\cos(x) - \sin(x)$. We have measured the correlations $C(t)$ and $K(t)$ as well as responses $\chi(t)$ for small h.

Fig. 6.2(a) shows that for strong stationary nonequilibrium $f \gg 0$ there is a large difference between the integrated correlation function of the traffic and that of the entropy production. However, their average $C_{NE}(t)$ agrees very well with the response $\chi(t)$, as it should. It is also important to remember that our approach works for non-stationary regimes as well. In Fig. 6.2(b) we show an example of a particle starting at time $t = 0$ from a non-stationary initial distribution $\mu(x) = \delta_{x,0}$ (i.e. its position is $x_0 = 0$), but for the rest not forced outside equilibrium, $f = 0$, to emphasize the transient character of this situation. Again, we can see that the response is well estimated by $C_{NE}(t)$.

In the next chapter we treat the inertial version of this model.

6.2.3 A generalized Einstein relation

Consider the overdamped Langevin equation in $\mathbb{R}^d$ (4.13) but with constant $\nu = \beta D$:

$$dx_t^i = \nu^{ij}[F^j(x_t) + h\delta^{jk}]dt + \sqrt{2D}^{ij} dB_t^j \tag{6.10}$$

with repeated indices $j = 1, \dots d$ summed over. The F includes all forces and h is the small constant perturbation. Note that this means that the perturbation potential is $V(x) = x^k$. With such a potential the fluctuation-dissipation relation connects mobility and diffusion which are actually defined as follows. The true mobility M (in contrast to ν) is defined as the response

$$M^{ik} = \lim_{t\to\infty} \frac{d}{dh} \left\langle \dot{x}_t^i \right\rangle_\rho^h \Big|_{h=0} = \lim_{t\to\infty} \frac{1}{t} \int_0^t \frac{d}{dh} \left\langle \dot{x}_t^i \right\rangle_\rho^h \Big|_{h=0} \tag{6.11}$$

while the real diffusion matrix (in contrast to D) is given by

$$\mathcal{D}^{ik} = \lim_{t\to\infty} \frac{1}{2t} \int_0^t ds \int_0^t dr \ \left\langle (\dot{x}_s^i - \left\langle \dot{x}^i \right\rangle_\rho)(\dot{x}_r^k - \left\langle \dot{x}^k \right\rangle_\rho) \right\rangle_\rho^0$$

As (6.11) is a response function, we can easily compute it within our framework. Note that $LV(x) = \nu^{kj} F^j(x)$:

$$\frac{d}{dh} \left\langle \dot{x}_t^i \right\rangle_\rho^h \Big|_{h=0} = \frac{\beta}{2} \left\langle [x_t^k - x_0^k] \dot{x}_t^i \right\rangle_\rho^0 - \frac{\beta \nu^{kj}}{2} \int_0^t ds \left\langle F^j(x_s) \dot{x}_t^i \right\rangle_\rho^0$$

We rewrite this response function by using the following equality, derived from taking the average of the Langevin equation ((6.10) with $h = 0$):

$$\left\langle \dot{x}^i \right\rangle_\rho^0 = \left\langle \nu^{ij} F^j(x) \right\rangle_\rho^0$$

because the average of dB_t gives zero. Using this equality we get

$$\begin{aligned} \frac{d}{dh} \left\langle \dot{x}_t^i \right\rangle_\rho^h \Big|_{h=0} &= \frac{\beta}{2} \int_0^t ds [\left\langle \dot{x}_s^k \dot{x}_t^i \right\rangle_\rho^0 - \left\langle \dot{x}_s^k \right\rangle_\rho^0 \left\langle \dot{x}_t^i \right\rangle_\rho^0] \\ &\quad - \frac{\beta \nu^{kj}}{2} \int_0^t ds \left\langle F^j(x_s)(\dot{x}_r^i - \left\langle \dot{x}^i \right\rangle_\rho^0) \right\rangle_\rho^0 \end{aligned}$$

Now all that is left is integrating this relation over t, dividing by t and taking the limit for $t \to \infty$ to arrive at

$$M^{ik} = \beta \mathcal{D}^{ik} - \frac{\beta \nu^{kj}}{2} \lim_{t\to\infty} \frac{1}{t} \int_0^t ds \int_0^t dr \left\langle F^j(x_s)(\dot{x}^i_r - \left\langle \dot{x}^i \right\rangle^0_\rho) \right\rangle^0_\rho \tag{6.12}$$

Observe that in equilibrium (F derives from a potential), the second term in (6.12) vanishes because the observable $\dot{x}_t$ is anti-symmetric, and $\langle \dot{x} \rangle^{\rm eq}_\rho = 0$. Then, the equilibrium fluctuation-dissipation relation holds with $M = \beta \mathcal{D}$. Moreover, when $F = 0$, (pure diffusion) these quantities can be explicitly calculated, and are found to be equal to $M^{ik} = \nu^{ik}, \mathcal{D}^{ik} = D^{ik}$. We see that for pure diffusion the fluctuation-dissipation theorem is equivalent to the Einstein relation. But even for equilibrium systems, when $F \neq 0$ the fluctuation-dissipation theorem is not equivalent with the Einstein relation, although there is still a simple relation between mobility and diffusion.

Out of equilibrium, this simple relation is violated, as there is a correction, namely the second term in (6.12)). Note that this is completely compatible with the condition $\nu = \beta D$, because that is actually a relation for the reservoir, which is in thermal equilibrium at inverse temperature β. Similar results for a one-dimensional overdamped diffusion were reported before in [8, 97].

6.3 Connection with effective temperature

A reason, not mentioned before, that the fluctuation-dissipation theorem in equilibrium is so useful and powerful is that it specifies the temperature as a universal parameter: regardless of what potential we use and what observable, the relation between response and correlation function is always governed by that same single parameter. In this sense it is not strange to wonder what remains of this parameter out of equilibrium.

Therefore, even though the fluctuation-dissipation theorem is violated out of equilibrium, one attempts to restore it by the introduction of an effective temperature $T^{\rm eff}$, in the sense

$$R^\mu_{QV}(t,s) = \frac{1}{k_B T^{\rm eff}} \frac{d}{ds} \langle V(s) Q(t) \rangle^0_\mu \tag{6.13}$$

Many studies have been devoted to the study of this prefactor, in what sense it perhaps resembles a thermodynamic temperature-like quantity for some classes of observables and over some timescales; we refer to [12, 17, 51, 60] for an entry into the extensive literature. Clearly, whatever the purpose of the discussion, an

exact expression of the response should help, especially when entirely in terms of explicit correlation functions. The first calculations in this sense are in [14] and they have been referred to as the "no field-method" [91]. In particular, for purposes of simulation or numerical verification of (6.13) we do no longer need to perform the perturbation by hand. In fact, now we can write the ratio $T/T^{\rm eff} = X$ entirely in terms of correlation functions

$$X = X_{QV}(\mu; t, s) = \frac{1}{2}\big[1 - \frac{\langle LV(s)\, Q(t)\rangle^0_\mu}{\frac{d}{ds}\langle V(s)Q(t)\rangle^0_\mu}\big] \tag{6.14}$$

with numerator and denominator in (6.14) each having a specific physical meaning in terms of entropy and traffic. An effective temperature is obtained as the ratio between the traffic and the entropic term: if for some observables (V, Q) and over some time-scales,

$$Y\ \frac{d}{ds}\langle V(s)\, Q(t)\rangle^0_\mu = \langle LV(s)\, Q(t)\rangle^0_\mu$$

for some Y, then $X = (1 - Y)/2$. Equilibrium has $X = 1 = -Y$. In the case where $LV \approx 0$ as for a conserved quantity, then $Y = 0$ and $T^{\rm eff} = 2T$. Finally, X and the effective temperature $T^{\rm eff}$ get negative when the traffic term overwhelms the entropic contribution.

One should understand that (6.13) represents a rather optimistic scenario. Formula (6.13) mimics (5.2) by replacing a function depending a priori on the observable, the potential, the initial distribution and on other parameters as temperature and time, by just one parameter. Why should there be also out-of-equilibrium a single parameter and a useful notion of temperature in its usual thermodynamic understanding? Moreover, how would it depend on the observables V and Q? (See [77] for a very recent discussion.) Answers to these questions have been partially given but are often restricted within a context of mean field systems or for small fluctuations, effectively dealing with calculations as in Section 6.3.2, similar to calculations for scalar fields as in [21] and in [12]. In fact, the optimism in (6.13) is a sort of conservatism as it tries to return to equilibrium-like formulae. We take a different attitude: the violation of the equilibrium fluctuation-dissipation relation (FDR) is an opportunity to discover new connections between response and the relevant newly emerging physical quantities as traffic.

6.3.1 Explicit calculations for pure diffusions

In simple overdamped diffusion equations, one can explicitly calculate the correlation functions, and see when the concept of effective temperature could

make sense. Similar calculations have been done in [12, 21]. The simplest example is Brownian motion in one dimension:

$$dx_t = h_t dt + \sqrt{2D} dB_t$$

where we have taken the potential $V(x)$ equal to the position x. The generator of the unperturbed dynamics is (see Appendix A) $L = D\Delta$, so that $LV = 0$ and $T^{\rm eff} = 2T$ cf. Virasoro's example in [21]. In that last reference, what are called "flat directions" can be associated to perturbations with $LV = 0$.

6.3.2 Explicit calculations for linear diffusions

Next, consider the following model for diffusions with a linear force (harmonic potential). Fix parameters $\alpha, B \in \mathbb{R}, D > 0$ and look at the linear Langevin dynamics for a global order parameter $M \in \mathbb{R}$,

$$dM(t) = -\alpha\, M(t)dt + h_t\, Bdt + \sqrt{2D}\, dB(t)$$

The $h_t, t > 0$, is a small time-dependent field. The generator of such a dynamics dynamics (on observables f) is (see Appendix A) $L^h f(M) = [-\alpha\, M + h_t B]\, f'(M) + D\, f''(M)$, using a prime to denote differentiation with respect to M. Such a dynamics can arise as a Gaussian approximation to a relaxational dynamics of the scalar magnetization M (no conservation laws and no spatial structure) valid in high enough dimensions (above $d = 4$ for the standard Ising model). Then, in a way, $\alpha = 0$ corresponds to the critical (massless) dynamics and $\alpha > 0$ is a paramagnetic dynamics (high temperature), see [12]. By taking $D \downarrow 0$ we exclude the diffusive aspects and we can think then of gradient relaxation in the low temperature regime.

The equilibrium (reversible stationary density on $\mathbb{R}$ for perturbation $B = 0$) is

$$\rho(M) = \frac{1}{Z} \exp\left\{-\alpha \frac{M^2}{2D}\right\}$$

with zero mean and variance $\langle M^2 \rangle = D/\alpha$. We now start from an initial distribution μ_0. If $\mu_0 \neq \rho$ then the system is not in equilibrium. We can compute the response function by calculating the expectation value of M at time t in the perturbed dynamics. We do that by deriving the following differential equation

$$\frac{d}{dt}\langle M(t)\rangle^h_{\mu_0} = \langle L^h M(t)\rangle^h_{\mu_0} = -\alpha \langle M(t)\rangle^h_{\mu_0} + h_t B$$

the solution of which is

$$\langle M(t)\rangle^h_{\mu_0} = \langle M(0)\rangle^0_{\mu_0}\, e^{-\alpha t} + B\int_0^t ds\, h_s\, e^{-\alpha(t-s)}$$

or

$$\frac{\delta}{\delta h_s}\,\langle M(t)\rangle^h_{\mu_0}\big|_{h=0} = B\, e^{-\alpha(t-s)} \tag{6.15}$$

which does in fact not depend on $\langle M(0)\rangle^0_{\mu_0}$ (and thus also equals the equilibrium result). In the same way, we use for $s < t$:

$$\frac{d}{dt}\langle M(t)M(s)\rangle^0_{\mu_0} = \langle M(s)L^0M(t)\rangle^0_{\mu_0} = -\alpha\langle M(t)M(s)\rangle^0_{\mu_0}$$

to arrive at

$$\langle M(t)M(s)\rangle^0_{\mu_0} = \langle M^2(s)\rangle^0_{\mu_0} e^{-\alpha t}$$

Using same strategy we can calculate the expectation of $M^2(s)$. Denoting $\overline{M_0^2} = \big\langle M^2(0)\big\rangle^0_{\mu_0}$, the correlation function for $0 < s < t$ is

$$\langle M(s)\, M(t)\rangle^0_{\mu_0} = \overline{M_0^2}\, e^{-\alpha(t+s)} + \frac{D}{\alpha}\big[e^{-\alpha(t-s)} - e^{-\alpha(t+s)}\big]$$

and hence

$$\frac{d}{ds}\langle M(s)\, M(t)\rangle^0_{\mu_0} = -\alpha\overline{M_0^2}\, e^{-\alpha(t+s)} + D\big[e^{-\alpha(t-s)} + e^{-\alpha(t+s)}\big]$$

If $\mu_0 = \rho$ in the last expression (thus replacing $\overline{M_0^2}$ by D/α) we find

$$\frac{d}{ds}\langle M(s)\, M(t)\rangle^0_{\rho} = D\, e^{-\alpha(t-s)}$$

which, in comparison with (6.15) specifies the equilibrium inverse temperature to be equal to $\beta = B/D$.

The traffic term is obtained from $LM = -\alpha M$, and thus

$$\langle LM(s)\, M(t)\rangle^0_{\mu_0} = -\alpha\, \overline{M_0^2}\, e^{-\alpha(t+s)} - D\big[e^{-\alpha(t-s)} - e^{-\alpha(t+s)}\big]$$

Clearly,

$$\frac{\delta}{\delta h_s}\langle M(t)\rangle^h_{\mu_0}(h=0) = \frac{B}{2D}\{\frac{d}{ds}\langle M(s)\,M(t)\rangle^0_{\mu_0} - \langle LM(s)\,M(t)\rangle^0_{\mu_0}\}$$

as it should.

For the issue of effective temperature we compute the ratio Y as

$$Y = Y(M_0;s,t) = \frac{-\alpha\langle M(s)M(t)\rangle^0_{\mu_0}}{\frac{d}{ds}\langle M(s)M(t)\rangle^0_{\mu_0}} = \frac{D-\alpha\overline{M_0^2}-De^{2\alpha s}}{D-\alpha\overline{M_0^2}+De^{2\alpha s}}$$

Remark that Y is independent of t, as long as $t > s$. In that notation, the effective inverse temperature is $T^{\rm eff} = 2T/(1-Y)$. In equilibrium $Y=-1$ while $Y=1$ for $D=0$ and $M_0 \neq 0$. If $\alpha = 0$, then $Y = 0$ as in the pure diffusion case. For $\alpha > 0$, if we let $s \to \infty$ while keeping $t-s=u$ fixed, we get

$$\lim_{s\uparrow+\infty} Y(M_0;s,s+u) = -1 \tag{6.16}$$

so that $T^{\rm eff} = T$. This case is referred to as the paramagnetic case, while $T^{\rm eff} = 2T$ for $\alpha = 0$ is the critical quench, cf. [12].

6.3.3 Effective traffic

Let us try to see how the notion of effective temperature could be seen as a one-parameter reduction of a general equilibrium-like FDR that is valid also outside equilibrium but with an effective dynamics. The starting point is comparing the equilibrium formula (5.2) with the more general non-equilibrium one (6.7). By this we see that in equilibrium (5.2) is equivalent with

$$R_{QV}(t,s) = -\beta\langle (LV)(s)\,Q(t)\rangle^0_\rho \tag{6.17}$$

In other words, the fluctuation-dissipation theorem in equilibrium can also be called a fluctuation-traffic theorem; the two terms on the right-hand side of (6.7) are simply the same in equilibrium. Therefore, for the purpose of getting closer to equilibrium response formulae one really has the choice to mimic either (5.2) or rather (6.17). The first leads to the ambition of effective temperature (6.13), the latter to the new notion of effective traffic. But the latter is also much richer. In fact, a combination of (6.7) with (A.6) shows that the exact nonequilibrium response formula (6.7) can indeed be written in the equilibrium form (6.17):

$$R^\mu_{QV}(t,s) = -\langle G_{\mu_s}V(s)Q(t)\rangle^0_\mu \tag{6.18}$$

with a new effective traffic

$$G_\mu V = \frac{\beta}{2\mu}[L^\dagger(\mu V) - VL^\dagger\mu + \mu LV] \tag{6.19}$$

Here $L^\dagger$ is the forward generator, defined by (see Appendix A, (A.2)):

$$\int dx g(x) Lf(x) = \int dx f(x) L^\dagger g(x) \qquad \forall f, g$$

where the integration has to be replaced by a sum for Markov jump processes. The operator G_μ acting on V in (6.19) has the following exact property: it is itself a generator but of a new dynamics for which μ is an equilibrium distribution (i.e. $\langle f\,(G_\mu g)\rangle^0_\mu = \langle g\,(G_\mu f)\rangle^0_\mu$ for all f, g, see (A.5)). Thus, in (6.18) the generator G_{μ_s} is the instantaneous equilibrium generator with respect to the time-evolved distribution μ_s.

We can rewrite G_μ in terms of the adjoint generator (see (A.4)), which is defined using the stationary distribution of the process:

$$\int dx \rho(x) g(x) L^* f(x) = \int dx \rho(x) f(x) Lg(x) \qquad \forall f, g$$

With this we get

$$G_\mu V = \frac{\beta\rho}{2\mu}[L^*(\frac{\mu}{\rho}V) - VL^*(\frac{\mu}{\rho}) + \frac{\mu}{\rho}LV] \tag{6.20}$$

In the stationary nonequilibrium case, we have $\mu_s = \rho$ and (6.19) is

$$G_\rho = \frac{\beta}{2}(L + L^*) \tag{6.21}$$

replacing L in (6.17). For equilibrium dynamics $L = L^*$ (see Appendix A), and (6.18) reduces to (6.17). In stationary nonequilibrium, if the perturbation V is 'time-direction independent' in the precise sense that $LV = L^*V$, then the nonequilibrium response reduces to the equilibrium formula (6.17) and $X = 1$. See [82] for very related conjectures and observations.

Chapter 7

Underdamped diffusions

In this chapter we treat the case of underdamped diffusions, for which the formula (6.7) is no longer correct. We can, however, still calculate the response function and express it in terms of correlation functions of observable quantities. For several explicit examples the response is calculated and simulated. This chapter describes work reported in [2].

7.1 Model

In this chapter we consider an extension of the one-dimensional underdamped diffusions explained in Section 4.1. There we considered the dynamics of just one particle. Now we consider k particles in d dimensions. This means that the state of the system is now given by $(x, v) = (x_1, x_2, \ldots, x_n; v_1, v_2, \ldots, v_n) \in \mathbb{R}^{2n}$ of positions and momenta, with $n = kd$. These particles are subject to mechanical forces: a potential $U(x)$ taking care of the coupling and pinning of the positions (the pinning is also thought to confine the positions to some finite volume), and a nonconservative forcing $f_i(x)$. Each particle is connected to its own heat bath at an inverse temperature β_i. This gives the following set of $2n$ Langevin equations:

$$\begin{aligned} dx_t^i &= v_t^i \, dt \\ m_i dv_t^i &= [f_i(x) - \frac{\partial U}{\partial x_i}(x_t) - m_i \gamma_i v_t^i] dt + \sqrt{2D_i} \, dB_i(t) \end{aligned} \tag{7.1}$$

Again, we make no distinction between upper or lower indices: $x_i = x^i$. To avoid confusion, we always use i, j, k to denote the components of the vectors, and s, t

to denote times. For the simple underdamped diffusions in (4.1), we found by the local detailed balance assumption the Einstein relation between diffusion and friction coefficient. Let us investigate what it gives here. For this we write down the path-probability density for paths $\omega = (x_t, v_t)_{0\le t\le T}$. In the same way as explained in Section 4.3 we find, formally:

$$\mathcal{P}_{x_0,v_0}(\omega) = \frac{1}{\mathcal{N}} \exp\left\{ -\sum_i \frac{1}{4D_i} \int_0^T dt[m_i \dot{v}_t^i - F_i(x_t) + m_i\gamma_i v_t^i]^2 \right\} \tag{7.2}$$

with $F_i = f_i - \frac{\partial U}{\partial x_i}$, and $\mathcal{N}$ a normalization factor. The entropy flux into the environment is then

$$S(\omega) = \frac{d\mathcal{P}_{x_0,v_0}}{d\mathcal{P}_{x_T,v_T}\theta}(\omega) = -\sum_i \frac{m_i\gamma_i}{D_i} \int_0^T dt v_i[m_i\dot{v}_t^i - F_i(x_t)] \tag{7.3}$$

Note that the quantity

$$m_i\dot{v}_i - F_i = -m_i\gamma_i v_i + \sqrt{2D_i}\frac{dB_i}{dt}$$

is the (stochastic) force on particle i originating from the i-th heat bath. The integral over time of this force times the velocity of the particle thus gives the heat flux from the i-th reservoir to the system. As a consequence $S(\omega)$ is the entropy flux into the environment if and only if $D_i = \frac{m_i\gamma_i}{\beta_i}$ for all i. Note that the system is driven out of equilibrium by the nonconservative force, but also by the difference of the temperatures of the heat baths β_i.

At time zero, the probability density of the system being in state (x, v) is denoted as always by $\mu_0(x, v)$. The Fokker-Planck equation determining the time-evolution of this probability in this case is

$$\frac{d}{dt}\mu_t + \nabla \cdot J_{\mu_t} = 0 \tag{7.4}$$

for $\nabla = (\nabla_x, \nabla_v)$ and for the probability current $J_\mu = (J_\mu^x, J_\mu^v)$ with

$$(J_\mu^x)_i = m_i v_i \mu, \qquad (J_\mu^v)_i = (f_i - \frac{\partial U}{\partial x_i})\mu - \gamma_i\, v_i\, \mu - D_i\, \frac{\partial \mu}{\partial v_i} \tag{7.5}$$

7.2 Perturbation

The perturbation is a potential $V(x)$ added to the unperturbed Hamiltonian $H_o = \sum_i \frac{m_i v_i^2}{2} + U(x) \to H_o - h_s\, V(x)$ with small time-dependent amplitude

$h_s, s \geq 0$. We compare the path-probabilities of the perturbed versus the unperturbed process, using the action. By (7.2) we find

$$\frac{d\mathcal{P}^h_{x_0,v_0}}{d\mathcal{P}^0_{x_0,v_0}}(\omega) = e^{-A(\omega)} = \exp\left\{\frac{S_{ex}(\omega) - \mathcal{T}_{ex}(\omega)}{2}\right\}$$

where the action is split in its time-antisymmetric and symmetric parts. The time-antisymmetric part is:

$$S_{ex}(\omega) = \sum_{i=1}^{n} \beta_i \int_0^t h_s \frac{\partial V}{\partial x_i}(x_s)\, v_s^i \, ds \tag{7.6}$$

Note the following: if all temperatures are the same, $\beta_i = \beta$, then $S_{ex}(\omega) = \beta \int_0^t h_s \frac{\partial V}{\partial s}(x_s) ds$, which is consistent with (5.10). The time-symmetric part of the action gives:

$$\mathcal{T}(\omega) \quad = \quad \sum_i \frac{1}{D_i} \int_0^t h_s \frac{\partial V}{\partial x_i}(x_s) \Big[\{f_i(x_s) - \frac{\partial U}{\partial x_i}(x_s)\} ds - m_i dv_s^i\Big] + o(h) \tag{7.7}$$

The last stochastic integral can be interpreted in the Itô or Stratonovitch sense, this does not matter in this case, see (4.19). However, the stochastic integral is the reason that we can't write the functional derivative of traffic as a state function. Indeed, dv_i contains information about two consecutive states, not one.

7.3 Result

The model that we use in this chapter is in contact with more than one heat bath, possibly at different temperatures. Therefore we can't use the formulae derived in Section 5.4. We therefore fall back on the more general formula (5.6). We consider general observables $Q(x_t, v_t)$ which we simply denote by $Q(t)$. The immediate application of (7.6) and (7.7) to (4.19) gives

$$\begin{aligned}\frac{\delta}{\delta h_s} \langle Q(t)\rangle^h_\mu\big|_{h=0} \quad &= \quad \sum_{i=1}^{n} \frac{\beta_i}{2} \langle \frac{\partial V}{\partial x_i}(x_s)\, v_s^i\, Q(t)\rangle_{\mu_0} \\ &\quad -\frac{1}{2}\sum_{i=1}^{n} \frac{1}{D_i} \langle \frac{\partial V}{\partial x_i}(x_s)\, [f_i(x_s) - \frac{\partial U}{\partial x_i}(x_s) - m_i \dot{v}_s^i]\, Q(t)\rangle_{\mu_0}\end{aligned}$$

However, because of the singular nature of the white noise, $\dot{v}$ is not a good observable. We therefore rewrite that part of the correlation function. We take the stochastic integral in the Itô sense, which is easiest here. This means that we have to interpret the correlation function containing $\dot{v}_i$ as follows:

$$\left\langle \frac{\partial V}{\partial x_i}(x_s)\, \dot{v}_s^i Q(t) \right\rangle_{\mu_0} = \lim_{\epsilon\to 0} \frac{1}{\epsilon} \left\{ \left\langle \frac{\partial V}{\partial x_i}(x_s)[v_{s+\epsilon}^i - v_s^i] Q(t) \right\rangle_{\mu_0} \right\}$$

We rewrite this inspired by the product rule for derivatives to

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \left\{ \left\langle [\frac{\partial V}{\partial x_i}(x_{s+\epsilon}) v_{s+\epsilon}^i - \frac{\partial V}{\partial x_i}(x_s) v_s^i] Q(t) \right\rangle_{\mu_0} \right\}$$

$$- \lim_{\epsilon\to 0} \frac{1}{\epsilon} \left\{ \left\langle [\frac{\partial V}{\partial x_i}(x_{s+\epsilon}) - \frac{\partial V}{\partial x_i}(x_s)] v_{s+\epsilon}^i Q(t) \right\rangle_{\mu_0} \right\}$$

Finally taking the limit, we obtain a useful form:

$$\begin{aligned} \left\langle \frac{\partial V}{\partial x_i}(x_s)\, \dot{v}_s^i Q(t) \right\rangle_{\mu_0} &= \frac{d}{ds} \langle \frac{\partial V}{\partial x_i}(x(s))\, v_s^i\, Q(t) \rangle_\mu \\ &\quad - \sum_j \langle \frac{\partial^2 V}{\partial x_j \partial x_i}(x_s)\, v_s^j\, v_s^i\, Q(t) \rangle_\mu \end{aligned} \tag{7.8}$$

With this the response function becomes

$$\begin{aligned} \frac{\delta}{\delta h_s} \langle Q(t) \rangle_{\mu_0}^h \big|_{h=0} &= \frac{1}{2} \sum_{i=1}^n \beta_i \langle \frac{\partial V}{\partial x_i}(x_s)\, v_s^i\, Q(t) \rangle_\mu \\ &\quad - \frac{1}{2} \sum_i \frac{1}{D_i} \langle \frac{\partial V}{\partial x_i}(x_s)\, [f_i(x_s) - \frac{\partial U}{\partial x_i}(x_s)]\, Q(t) \rangle_{\mu_0} \\ &\quad + \frac{1}{2} \frac{d}{ds} \sum_i \frac{m_i}{D_i} \langle \frac{\partial V}{\partial x_i}(x_s)\, v_s^i\, Q(t) \rangle_{\mu_0} \\ &\quad - \frac{1}{2} \sum_{i,j} \frac{m_i}{D_i} \langle \frac{\partial^2 V}{\partial x_j \partial x_i}(x_s)\, v_s^j\, v_s^i\, Q(t) \rangle_{\mu_0} \end{aligned} \tag{7.9}$$

This may be a complicated formula, but we now have expressed the response function in terms of correlation functions of observables which are in principle measurable.

7.4 Examples

The result (7.9) is a complicated formula. This is why we have investigated it in several physically interesting examples. For two of these examples, simulations have been provided by Marco Baiesi [2].

We remind that a regime out of equilibrium can be created in (7.1) by letting the inverse temperatures β_i differ from one another. Another way of going out of equilibrium is to introduce nonconservative forces $\{f_i\}$, like external fields that are rotational. A final possibility that we can consider is to start from an initial condition that is not stationary, and thus to observe the response in a transient regime. All of these possibilities are covered in the following examples. As in the previous chapter, numerical results are better presented with integrated responses. We use slightly different definitions than in Section 6.2:

- The integrated response, also called generalized susceptibility:

$$\chi(t) = \int_0^t ds\, R_{QV}(t,s) = \lim_{h\to 0} \frac{\langle Q(t)\rangle_{\mu_0}^h - \langle Q(t)\rangle_{\mu_0}^0}{h}$$

- The correlation with entropy:

$$C(t) = \frac{1}{h} \langle S(\omega) Q(t)\rangle_{\mu_0}$$

- The correlation with traffic:

$$K(t) = \left\langle \frac{\partial}{\partial h}\, \mathcal{T}(\omega)|_{h=0}\, Q(t) \right\rangle_{\mu_0}$$

- The average

$$C_{\rm ne}(t) = \frac{C(t) + K(t)}{2}$$

Contrary to the examples in the last chapter, here we have embedded β_i's in the definitions of the correlation functions. The integrated response relation is thus $\chi(t) = C_{\rm ne}(t)$. We choose for simplicity particles with mass equal to one in all the examples.

7.4.1 Langevin particle in a periodic potential

Recently there have been experiments testing the response of an overdamped particle (high viscosity limit) in a periodic potential [45]. In the previous chapter,

section 6.2.2, we have discussed simulations of that system. Here we look for the changes in an underdamped set-up, allowing e.g. for the particle to have a considerable mass and to obey a noisy Hamiltonian dynamics.

We consider here a system consisting of one particle in one dimension: its position is $x_t \in S^1$ (on a circle) and the velocity is $v_t \in \mathbb{R}$. Then, the equations (7.1) simplify to

$$\begin{aligned} dx_t &= v_t dt \\ dv_t &= F(t)dt - \gamma v_t dt - h_t\, g(t)dt + \sqrt{2D}\, dB_t \end{aligned} \tag{7.10}$$

where we abbreviate

$$\begin{aligned} F(t) &= f - \frac{dU}{dx}(x_t) \qquad \text{(deterministic force)} \\ g(t) &= -\frac{dV}{dx}(x_t) \qquad \text{(perturbing force)} \end{aligned}$$

The nonconservative force f is the driving and is taken constant over the circle, thus effectively tilting the conservative potential. This drives the system out of equilibrium. As there is only one particle, there is only one heat bath.

At time $s = 0$ the unperturbed system is taken to be in the stationary nonequilibrium ρ corresponding to (7.4), and then for $s > 0$ a constant small h is turned on. Hence, the integrated correlations are

$$\begin{aligned} C(t) &= = -\beta \int_0^t ds\ \langle v_s g(s) Q(t)\rangle_\rho^0 = \beta\,[\langle V(t)Q(t)\rangle_{\mu_0}^0 - \langle V(0)Q(t)\rangle_{\mu_0}^0] \\ K(t) &= = \frac{1}{D}\left\{\int_0^t ds\,\langle F(s)g(s)Q(t)\rangle_\rho^0 - \int_0^t \langle dv(s)g(s)Q(t)\rangle_\rho^0\right\} \end{aligned} \tag{7.11}$$

Note that we left dv_s in the definition of $K(t)$. This can of course be rewritten as explained in Section 7.3. However, in simulations one discretizes time, and dv_s is known at every time-step. It can therefore be used as an observable here. We take $V(x) = U(x) = \cos(2\pi x)$ like in the previous chapter, and again also $Q = U$.

In Fig. 7.1 we visualize the various terms of the fluctuation-dissipation relation, for three scenarios with different viscosity, increasing from left to right. The response is well reproduced by $C_{\rm ne}(t)$, even if we perform a numerical integration with $dt = 10^{-3}$. Oscillations in the response are visible for small viscosity; at higher

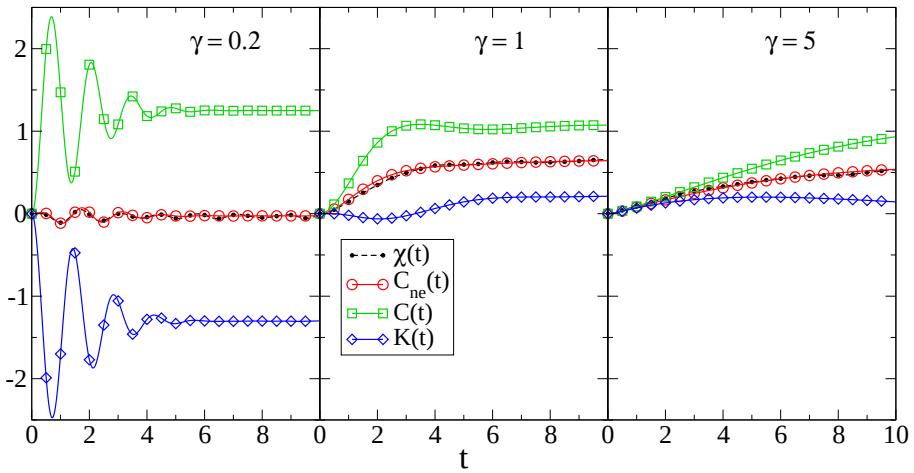


Figure 7.1: Integrated correlation functions of the entropic term $C(t)$, of the traffic term $K(t)$, and their average $C_{\rm ne}(t)$ giving the response in nonequilibrium, and the integrated response $\chi(t)$ calculated directly with $h = 0.01$. Panels are for simulations with various friction coefficients: $\gamma = 0.2$ (left), $\gamma = 1$ (center), and $\gamma = 5$ (right). Other parameters: $T = 1/\beta = 0.2$, $f = 0.9$.

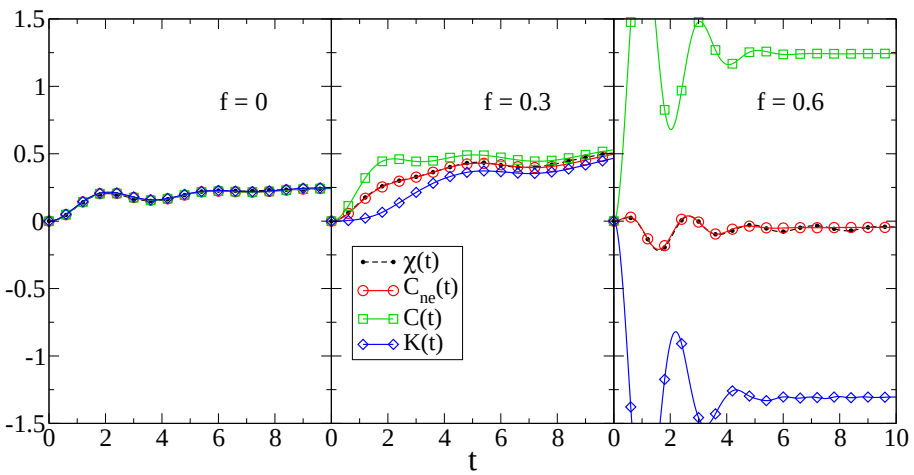


Figure 7.2: As in figure 7.1, but with fixed $\gamma = 0.2$ and varying force f, from left to right: $f = 0$ (equilibrium), $f = 0.3$, and $f = 0.6$ (case $f = 0.9$ is in the previous figure).

friction there is a monotonous drift towards a new stationary state (right panel in Fig. 7.1). The inertial regime is less sensitive to the perturbation, as $\chi(t)$ displays only a small wiggling: a high entropy production is almost compensated by a high traffic. For the high viscosity regime, with this setting ($f \lesssim 1$) the traffic term is close to zero compared with the entropic term, and their combination yields $C_{\rm ne} \approx C/2$.

In Fig. 7.2 we can follow the response as a function of the driving f, with $f = 0$ for equilibrium. For $f \neq 0$ the entropic term $C(t)$ can be quite different from $C_{\rm ne}(t)$. In equilibrium all the terms coincide, as expected.

7.4.2 A generalized Einstein relation

Let us consider the underdamped version of the example in the last chapter, Section 6.2.3. For this we take the general Markov dynamics (7.1), with the following simplifications: we take $V(x) = x_j$ for an arbitrary j, and for the observable we take $Q = v_k$ for a k which can be different from j. We also take the field $h_t = h$ to be constant. In this case the response function is related to the mobility of the system, i.e. the way in which the average velocity changes under a constant added force. We fix a large time u and define the time-averaged mobility by:

$$\mathcal{X}_{jk} = \frac{1}{u}\int_0^u dt\, \frac{\partial}{\partial h} \left\langle v_t^k \right\rangle_{\mu_0}^h \Big|_{h=0}$$

starting from μ_0 at time zero. We wish to connect this to the velocity fluctuations in the unperturbed (but driven) system:

$$\mathcal{D}_{jk} = \frac{1}{2u}\left\langle [x_u^j - x_0^j][x_u^k - x_0^k] \right\rangle_{\mu_0}^0 = \frac{1}{2u}\int_0^u dt \int_0^u ds \left\langle v_s^j v_t^k \right\rangle_{\mu_0}^0$$

We try to be more general here than in the last chapter: we do not take the limit $u \uparrow +\infty$. In equilibrium, when μ_0 is the Maxwell-distribution and with f and U equal to zero and all the temperatures are equal ($\beta_i = \beta$), one can compute that $\mathcal{X}_{jk} = \frac{1}{\gamma_j}\delta_{j,k}$ and $\mathcal{D}_{jk} = \frac{1}{\beta\gamma_j}\delta_{j,k}$. Indeed, also in the underdamped case, the Einstein relation coincides with the fluctuation-dissipation relation for pure diffusion. For general equilibrium systems we still have $\mathcal{X}_{jk} = \beta\mathcal{D}_{jk}$. This is no longer true out of equilibrium. With (7.9) we can give the explicit modification.

We substitute $V = x_j$ and $Q = v_k$ into the integrated version of (7.9) for constant h to obtain

$$\frac{\partial}{\partial h} \left\langle v_t^k \right\rangle_{\mu_0}^h \Big|_{h=0} \quad = \quad \frac{1}{2}\beta_j \int_0^t ds \left\langle v_s^j v_t^k \right\rangle_{\mu_0}$$

$$-\frac{1}{2D_j}\int_0^t ds\left\langle [f_j(x_s)-\frac{\partial U}{\partial x_j}(x_s)]v_t^k\right\rangle_{\mu_0}$$

$$+\frac{1}{2D_j}\int_0^t ds\frac{\partial}{\partial s}\left\langle v_j^s v_t^k\right\rangle_{\mu_0}$$

The integrand of the full right-hand side is zero for $s > t$ due to causality, so we can as well integrate s from 0 up to $u > t$. Integrating then t gives us

$$\begin{aligned}\mathcal{X}_{jk} &= \beta_j\mathcal{D}_{jk}-\frac{1}{2uD_j}\int_0^u dt\int_0^u ds\left\langle [f_j(x_s)-\frac{\partial U}{\partial x_j}(x_s)]v_t^k\right\rangle_{\mu_0}\\ &\quad+\frac{1}{2uD_j}\int_0^u dt\left\langle [v_t^j-v_0^j]v_t^k\right\rangle_{\mu_0}\end{aligned}$$

This relation reduces to the familiar Einstein relation if the unperturbed system is in equilibrium. Furthermore very formally when the overdamped limit is taken, i.e. we neglect changes in momenta, then the last term in the relation drops and we recover the relation found in (6.12).

7.4.3 Coupled oscillators

We now consider coupled one-dimensional oscillators at different temperatures T_i in the stationary or in a transient regime. This means that each index i represents a particle which can move in one dimension. In (7.1) we then set a conservative potential U that is the sum $U = \sum_{i=0}^{n}\varphi(x_{i+1}-x_i)$ of local couplings between the oscillators i and $i+1$, with $\varphi(x) = \frac{1}{2}x^2+\frac{1}{4}x^4$. Boundary conditions are imposed by keeping $x_0 = x_{n+1} = 0$. For simplicity we take all $D_i = D$. A basic perturbation is given by switching on an external field on the j-th particle $V(x) = -Ex_j$. Taking the velocity $Q = v_k$ at site k to be the observable, the variable excess in entropy flux (7.6) reduces to

$$S_{ex}(\omega) = -\beta_j E\int_0^t ds\, v_s^j\, h_s$$

while the traffic equals

$$\mathcal{T}(\omega) = -\frac{E}{D}\int_0^t ds\left[\dot{v}_s^j+\frac{\partial U}{\partial v_j}(x_s)\right]h_s$$

From (7.9) we have

$$\frac{\delta}{\delta h_s}\left\langle v_t^k\right\rangle_\rho^h\Big|_{h=0} = -\beta_j E\left\langle v_s^j v_t^k\right\rangle_\rho^0 - \frac{E}{2D}\frac{d}{ds}\left\langle v_s^j v_t^k\right\rangle_\rho^0 - \frac{E}{2D}\left\langle \frac{\partial U}{\partial x_j}(x_s)\, v_t^k\right\rangle_\rho^0 \tag{7.12}$$

This last relation is still valid for all times s, t and is automatically equal to zero for $s > t$ (causality). Remembering that $\beta_j = \frac{\gamma_j}{D} = 1/T_j$ we can rearrange formula (7.12) for the situation where $s < t$ (like we did in (5.16)):

$$\frac{d}{dh_s}\left\langle v_t^k\right\rangle_\rho^h = -E\left(\frac{\beta_j+\beta_k}{2}\right)\left\langle v_s^j v_t^k\right\rangle_\rho^0 - \frac{E}{2D}\left(\left\langle \frac{\partial U}{\partial x_j}(x_s)\, v_t^k\right\rangle_\rho^0 + \left\langle v_s^j \frac{\partial U}{\partial x_k} x_t\right\rangle_\rho\right)$$

Note that the right-hand side now shows a formal space-time symmetry for exchanging $j \leftrightarrow k, s \leftrightarrow t$. In equilibrium the symmetry is true on spatial level alone, $j \leftrightarrow k$, because time-symmetry is automatic. That is then an instance of Onsager reciprocity [85, 86].

Choosing constant $h = 0.03$, in Fig. 7.3(a) we show the response in a system with $n = 11$ oscillators, with a linear gradient of temperature $T_i = i/10 = 1/\beta_i$, perturbation applied on site $j = 1$ and response tested at central site $k = 6$. In Fig. 7.3(b) instead we have constant temperatures $T_i = 0.2$, but we start from a state out of equilibrium, by choosing $p_1(0) = 10$ and the other momenta equal to zero.

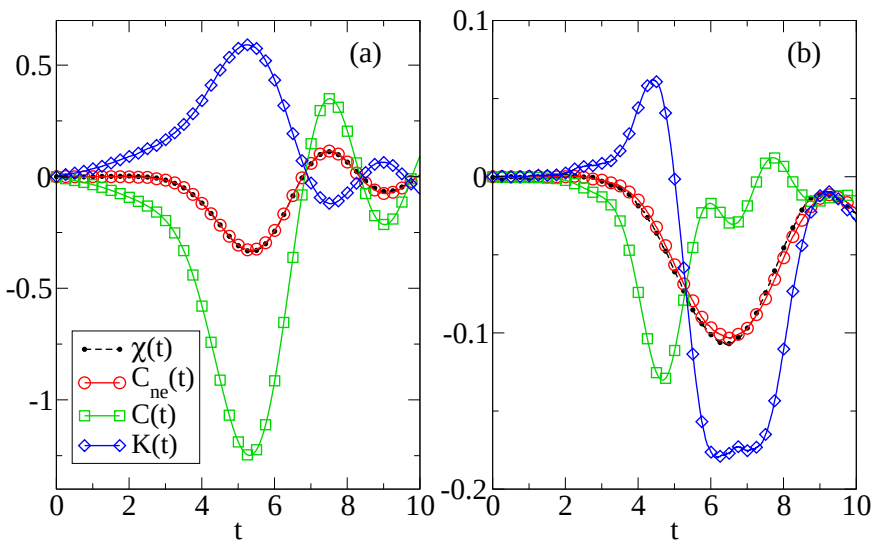


Figure 7.3: Visualization of the fluctuation-response relation $\chi(t) = C_{\rm ne}(t)$, for $n = 11$ coupled oscillators, with parameters, $D = 0.01$, $h = 0.03$, and (a) $T_i = i/10$, (b) $T_i = 0.2$ and transient regime as described in the text.

Chapter 8

Conclusions and literature

To conclude this part of the thesis we give an overview of previous formulations of nonequilibrium fluctuation dissipation relations, and present our conclusions of our own work.

8.1 Overview of previous formulations

The literature on (extensions of) the fluctuation-dissipation theorem is vast, and we can't list all possible and even essential contributions. The equilibrium formulation spans all of the previous century, while nonequilibrium versions started to appear since the 1970's and very much continue up to now. Early works include [1, 102] and also [23, 47], where we see a discussion within the theory of stochastic dynamics. In contrast to [47, 48] our unperturbed process is time-homogeneous. Coming to more recent times, violations of the equilibrium relation have been most often discussed in transient regimes (equilibrium dynamics but not relaxed to equilibrium). For example, in the context of ageing phenomena [17, 51], much thought has been given to making sense of an effective temperature as was discussed in Section 6.3.

However, recently much new work has been also directed to find generic extensions of the fluctuation-dissipation theorem to nonequilibrium steady states and to discussions of the dissipative elements in relaxations to nonequilibrium. There is for example the line starting from Sasa *et al* [49, 50, 82, 88] which treats nonequilibrium heat effects. In particular [49] might also be useful in real experiments because the response function is there directly connected with the energy dissipation, which we also discussed in Section 5.4.3. We do not know yet how to relate that to the new ideas surrounding the traffic term in our work.

For other recent extensions of the FDR, we refer to [6, 8, 40, 97, 98]. We have also mentioned in [2, 4] (but not in this thesis) how our approach is related to the co-moving frame interpretation of Chetrite, Gawędzki *et al* [15, 16, 45]. There the equilibrium fluctuation-dissipation relation is recovered for systems out-of-equilibrium by going to a description in a co-moving frame. One disadvantage of this approach, and of the approaches in the previous references, is that one keeps the stationary density ρ (or its logarithm) as an observable in the correlation functions. In our approach the largely unknown distribution only enters in the statistical averaging. On the other hand, the other approaches come with a new and still interesting interpretation (e.g. response in the time-reversed dynamics in [6] and co-moving frame in [15, 16]).

A first generalized fluctuation-dissipation relation giving a response formula *in our sense* appears in Section 2 of [21] (Kurchan *et al*). It treats an overdamped Langevin dynamics for a variable $y(t)$, used for soft spin models, and the resulting equation (2.10) in [21] is exactly our formula (5.16) for $Q = V$ equal to y, where (6.6) can be used. Another treatment for systems of Ising spins is done in [63, 64] (Lippiello *et al*), there the response is also investigated to higher orders in h. A more general treatment for jump processes is offered in [26] (Diezemann), in particular its equations (16)-(17), see also [75]. In contrast, we have emphasized the interpretation via entropy and traffic in nonequilibrium fluctuation theory (more is to come on the meaning of traffic in the next part of this thesis). In fact, that interpretation is exactly what makes a systematic generalization possible at all. The only study in which we recognize some of the ideas related to the traffic term is in [93] (Ruelle), in the context of dynamical systems.

8.2 Conclusions

We have studied linear response relations under general nonequilibrium conditions (stationary and not). From physical constraints on the probability of trajectories, we have obtained in Chapter 5 a general fluctuation-dissipation relation for the response of a driven system to the addition of a potential. Most generally in formula (5.6), and more specifically for systems in contact with an environment at one specific temperature in (5.14). Our scheme lifts the nonequilibrium FDR beyond formal first order perturbation theory applied to a specific dynamics, by identifying in general physical terms the statistical quantities that determine the response: entropy and traffic.

Chapter 6 has dealt with jump processes and with overdamped diffusions, which is the usual set-up for discussions on the violation of the fluctuation-dissipation theorem and for the possible emergence of an effective temperature. The general fluctuation-dissipation relation for these classes of models is (6.7). As in this relation the correlations with entropy and with traffic are expressed in terms of

explicit averages, they constitute a formula ready to use in a general context. For example, estimates of these correlations can be obtained with usual averaging in simulations, without any need to know further details or approximations of the stationary density of states. On the theoretical side, several previous approaches can be recovered or extended within the our scheme. In many cases these results have been discussed in specific model dynamics or for specific observables. It is interesting that there is a unifying approach wìth statistical interpretation behind this very broad variety of previous results.

In Chapter 7, we have applied our scheme to underdamped diffusions, resulting in (7.9). It is seen that the correlation with entropy involves dissipation over different reservoirs, each one with its own equilibrium temperature. The correlation with traffic is again written in terms of observable quantities.

Generally, we like to stress the emergence of traffic, as an important player out of equilibrium, complementary to the entropy flux. In equilibrium entropy and traffic give the same contribution to the response function. Out of equilibrium, they detach and the response needs to be evaluated in terms of correlation functions with both entropy and traffic. So far traffic is in general not understood operationally for real experiments, but in the previous chapters it has been expressed as a statistical average and correlation function of theoretically known forces. The fluctuation-dissipation relations we derived also give observational significance to the notion of traffic, which so far has mostly appeared as a theoretical concept in fluctuation theories (see also the next part of this thesis).

In the ongoing research it will be interesting to consider more general and other models, e.g even non-Markovian systems. Also within the models we considered, more general types of perturbation are possible and interesting. For example, what will happen if the perturbation is not of the potential type but actually changes the nonequilibrium part of the dynamics? Think of changing one of the temperatures of several heat baths, or adding an extra nonconservative forcing. In all these cases our results will have to be modified, of course, but we believe that the general method and framework we use is applicable to this wider set of questions. Specifically, the concepts of entropy and traffic will remain major actors in nonequilibrium systems.

Part III

Dynamical fluctuations

"Large deviations estimates have proved to be the crucial tool required to handle many questions in statistics, engineering, statistical mechanics, and applied probability."

Amir Dembo and Ofer Zeitouni, in their preface to *Large Deviations Techniques and Applications* (1998), [24]

Chapter 9

Large deviations in statistical mechanics

In 1910, Einstein proposed to invert the equation on Boltzmann's gravestone, to express probabilities of fluctuations in terms of entropy:

$$W = \exp\{S/k_B\}$$

Such a formula shows that the probability of a fluctuation (a deviation of the macrostates from their equilibrium values) is exponentially small and the entropy is the function that governs this exponential behaviour (rate function). This idea is generalized in the application of large deviation theory to statistical mechanics. This theory mathematically deals with the exponential decay of probabilities of large deviations in stochastic processes, as a correction to the law of large numbers. It is very useful and forms a natural mathematical formalism for statistical mechanics, because thermodynamic potentials as entropy and free energy naturally emerge as rate functions. This chapter forms a brief introduction for the next chapters.

9.1 Introductory example: a discrete ideal gas

Consider a discrete ideal gas [37]. We mean by this a system consisting of N independent particles, labelled by $i = 1, \ldots, N$. The energy ω_i of each particle is a stochastic variable, drawn from the set $\{E_k | k = 1, \ldots, m\}$. The probability that particle i has an energy $\omega_i = E_k$ is ρ_k, where $\rho = (\rho_1, \ldots, \rho_m)$ is a probability vector, i.e. $\sum_k \rho_k = 1$. The total state of the system is given by $\omega = (\omega_1, \ldots, \omega_N)$.

Because the particles are assumed to be independent, the probability of the whole state ω is then

$$\text{Prob}(\omega) = \prod_{i=1}^{N} \text{Prob}(\omega_i) = \prod_{i=1}^{N} \left[\sum_{k=1}^{m} \delta_{\omega_i, E_k} \rho_k \right]$$

As a first step towards a more macroscopic description of such a system, we are not interested in the exact energy of each particle, but rather in how many particles have a certain energy. We define therefore the fraction of particles that have an energy E_k:

$$p_\omega(k) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\omega_i, E_k} \tag{9.1}$$

where δ is the Kronecker delta. We call this fraction the empirical occupation of E_k. The empirical occupation vector for the collection of energies is

$$p_\omega = [p_\omega(1), \ldots, p_\omega(m)]$$

Note that this (macrostate) is a stochastic variable, and at the same time defines a probability distribution, because $\sum_k p_\omega(k) = 1$.

The law of large numbers By the law of large numbers, we expect this empirical occupation vector to resemble $\rho = (\rho_1, \ldots, \rho_m)$ when the particle number N is very large. More precisely, we define the space of probability vectors

$$\mathcal{X} = \{\mu = (\mu_1, \ldots, \mu_m) \in [0,1]^m | \sum_{k=1}^{m} \mu_k = 1\}$$

The collection of possible empirical occupation vectors for N particles form a subspace of $\mathcal{X}$, which we denote by Λ_N. The metric we consider on $\mathcal{X}$ is

$$d(\mu, \nu) = \sum_k (\mu_k - \nu_k)^2 \qquad \mu, \nu \in \Omega$$

Then the law of large numbers dictates that for any positive $\epsilon \in \mathbb{R}$

$$\lim_{N \to \infty} \text{Prob}(d(p_\omega, \rho) < \epsilon) = 1$$

However, for finite N, fluctuations/deviations from this expected behaviour are always possible. If we want to know more about these fluctuations, we have to go beyond the law of large numbers.

Central limit theorem If we want to go beyond the law of large numbers, we could use the central limit theorem. Formally, the central limit theorem then says that in a distributional sense

$$\text{Prob}(\frac{Np_\omega - N\rho}{\sqrt{N}} \sim z) \to \frac{1}{\mathcal{N}} \exp\{-\sum_{k=1}^{m} \frac{z_k^2}{2\rho_k}\} \tag{9.2}$$

for $z \in \mathbb{R}^m$ and $\mathcal{N}$ a normalization factor. To derive (9.2) is not trivial and we will not elaborate on this, because it is not very important for our purposes. We do stress that for large N this theorem only says something about the probabilities of events for which $p_\omega - \rho$ is of the order of $1/\sqrt{N}$). We call these small fluctuations. To know more, we have to consider the probabilities of large fluctuations.

Beyond the central limit theorem In a simple system as this, it is not difficult to compute the probabilities of arbitrary fluctuations. Let us therefore take an arbitrary $\mu \in \Lambda_N$. This μ can be totally different from ρ and is therefore called a (large) fluctuation. Elementary combinatorics give us

$$\text{Prob}(p_\omega = \mu) = \frac{N!}{(N\mu_1)! \cdot \ldots \cdot (N\mu_m)!} \rho_1^{\mu_1} \cdot \ldots \cdot \rho_m^{\mu_m}$$

For systems with large N this expression as such is not very convenient. We therefore use Stirling's approximation for large N to derive the following:

$$\frac{1}{N} \log \text{Prob}(p_\omega = \mu) \approx -\sum_{k=1}^{m} \mu_k \log\left(\frac{\mu_k}{\rho_k}\right) \tag{9.3}$$

The approximation becomes exact in the limit of $N \to \infty$. For this reason, we write the probabilities in the following form:

$$\text{Prob}(p_\omega = \mu) \approx e^{-NI(\mu)}$$

where $I(\mu) = \sum_{k=1}^{m} \mu_k \log \frac{\mu_k}{\rho_k}$ is minus the relative entropy between the probability vectors μ and ρ. One can easily prove that $I(\mu) \geq 0$ in general. Moreover, it is only zero for $\mu = \rho$.

Even though we have made an approximation, this result still gives more information than the law of large numbers and even more than the central limit theorem: it tells us that the probability that the empirical occupation vector to be equal to some $\mu \neq \rho$ is exponentially small in the number of particles (only when $\mu = \rho$ there is no exponential decay, in accordance with the law of

large numbers). At least, this is the dominant behaviour for large N. It is not very surprising that there is an exponential behaviour in the number of particles, because all particles are independent. However, also in many systems which do have interactions, this exponential behaviour is observed. This may not come as a surprise if we realize that many interactions are short-ranged, and parts of the system that are sufficiently far apart are still approximately independent.

Such an exponential behaviour as in (9.3) is an instance of a large deviation principle. As it turns out large deviation theory forms a natural framework in which statistical mechanics can be embedded. It incorporates in a natural way many key features, as the framework of thermodynamic potentials and the variational principles that characterize equilibrium.

In this chapter we introduce the reader into some aspects of the theory of large deviations, by using the example introduced above. The introduction that follows here is based on [37, 100], and for more explanation we therefore refer to those texts. For a thorough treatment of the mathematical theory, we refer to [24], while for the embedding of the theory of statistical mechanics in the large deviation formalism we refer to [35, 36, 61, 100].

9.2 The large deviation principle

In the example above, we defined the empirical occupation vector p_ω, and computed approximately its probability distribution for large N. Let us try to be mathematically more precise. We consider the probability that the empirical occupation vector is close to some given probability vector $\mu \in \mathcal{X}$. We thus take an arbitrary (small) positive $\epsilon \in \mathbb{R}$ and consider the probability $P(d(p_\omega, \mu) < \epsilon)$. Writing $B_N(\mu, \epsilon) = \{\nu \in \Lambda_N | d(\mu, \nu) < \epsilon\}$, this probability becomes

$$P(d(p_\omega, \mu) < \epsilon) = \sum_{\nu \in B_N(\mu,\epsilon)} P(p_\omega = \nu)$$

Although it is not trivial, one can prove [37] that this probability satisfies a large deviation principle, stated here in a simplified way:

$$-\lim_{N\to\infty} \frac{1}{N} \log P(d(p_\omega, \mu) < \epsilon) = \sum_{k=1}^{m} \mu_k \log\left(\frac{\mu_k}{\rho_k}\right) + O(\epsilon) = I(\mu) + O(\epsilon)$$

As this can be done for arbitrary small ϵ, one often denotes this by

$$-\lim_{N\to\infty} \frac{1}{N} \log P(p_\omega = \mu) = I(\mu) \tag{9.4}$$

The functional I is here the relative entropy between the probability vectors μ and ρ. Generally, in large deviation theory, I is called the rate function, as it quantifies the rate of exponential decay, or fluctuation functional., as it quantifies the probability of a fluctuation (deviation from the typical value).

To define the large deviation principle in general, one needs to go into many mathematical details [24]. This is not our goal. For our purposes, a definition as in (9.4) is enough. Moreover, we do not concern ourselves in this text with proving the large deviation principle. We do assume that, for the systems considered in the next chapters, those proofs can be done (in a physically uninteresting way). Instead we usually just assume the large deviation principle to hold, and refer to e.g. [24, 27] for full mathematical proofs.

Following [36, 100], we denote

$$P(p_\omega = \mu) \asymp e^{-NI(\mu)}$$

to express a large deviation principle as in (9.4).

We state here some properties of rate functions, without giving proofs, (for proofs we refer again to [100, 24]):

- For some $\mu \in \mathcal{X}$, the limit in (9.4) may not exist, because the probability density decays to zero faster than exponentially (super-exponentially) in N. In that case the rate function is set to $+\infty$.
- In some cases $I(\mu) = 0$ for some μ, meaning that the probability density decays slower than exponentially (sub-exponentially). In the example, this only happens when $\mu = \rho$. In the systems we consider in the next chapters, we assume that the rate function only has one global minimum where I is zero. Actually this is equivalent with the assumption that there is a unique stationary state of the system.
- Rate functions are positive: $I(\mu) \geq 0$. If they were not, the corresponding probabilities would diverge instead of decay.

9.3 The contraction principle

Suppose we are given a new random variable, defined through a continuous function of the old one: $q_\omega = h(p_\omega)$. As we will see in the following, such a

function can be many-to-one. A heuristic argument gives the rate function for q in terms of the rate function for p. Through the large deviation principle we write

$$p(q_\omega = \nu) \asymp \int_{\mu:h(\mu)=\nu} d\mu e^{-NI(\mu)}$$

Laplace's saddle-point approximation gives an exact expression for the following limit:

$$-\lim_{N\to\infty} \frac{1}{N} \log \int_{\mu:h(\mu)=\nu} d\mu e^{-NI(\mu)} = \inf_{\mu:h(\mu)=\nu} I(\mu)$$

One concludes that

$$p(q_\omega = \nu) \asymp \exp\{-N \inf_{\mu:h(\mu)=\nu} I(\mu)\}$$

defining the rate function for q_ω as $I_q(\nu) = \inf_{\mu:h(\mu)=\nu} I(\mu)$. This way of computing one rate function as the infimum of another is called a contraction.

Let us consider the following case. In absence of information about our system, we assume that the probabilities of all the energy levels of the particles are the same, i.e. $\rho_k = 1/m$. The rate function $I(\mu)$ then becomes $I(\mu) = \log m + \sum_k \mu_k \log \mu_k$. Up to some constants this is minus the Shannon entropy of the probability vector μ. Consider the probability that the total energy per particle (average energy) is equal to a value E, giving the following contraction:

$$I(E) = \inf_{\mu:\sum_k \mu_k E_k = E} I(\mu)$$

We can find such an infimum by using Euler-Lagrange equations with Lagrange multipliers to ensure the restrictions. We get k equations of the form:

$$0 = \frac{\partial}{\partial \mu_k}[\log m + \sum_k \mu_k \log \mu_k + \beta(\sum_k \mu_k E_k - E) + \gamma(\sum_k \mu_k - 1)]$$

where γ makes sure that μ sums to one and β makes sure that the average energy is E. The solution of these equations is then a μ of the form

$$\mu_k = \frac{1}{Z} e^{-\beta E_k}$$

where Z is a normalization constant, and β is determined by $\sum_k \mu_k E_k = E$. This is the same form as the probability distribution of the canonical ensemble! It arises

here as the most probable probability vector given an average energy E. The rate function becomes

$$I(E) = \log m - \beta E - \log Z$$

One can check that $\frac{dI}{dE} = \beta$. If we return to the idea that entropy and probability are related by $S = k_B \log P$, then the rate function I should represent minus the entropy per particle, divided by k_B. As a result $dS/N = k_B \beta dE$. Indeed, the Lagrange multiplier β is nothing but the inverse temperature associated to this system. We also see that $-\frac{1}{\beta} \log Z$ is the free energy of the system.

In this example, we clearly see that the ensemble theory of equilibrium statistical mechanics naturally appears in large deviation theory, at least for this example. The entropy plays the role of rate function, and the canonical ensemble is found through a contraction principle as the most probable probability vector given an average energy E.

9.4 The Gärtner-Ellis theorem

Let us take an arbitrary vector $M \in \mathbb{R}^m$. We denote the scalar product between such a vector and the probability vector μ by $M \cdot \mu = \sum_k M_k \mu_k$. The scaled cumulant generating function $\lambda(M)$ for our example is defined by

$$\lambda(M) = \lim_{N \to \infty} \frac{1}{N} \log \left\langle e^{NM \cdot p_\omega} \right\rangle$$

According to the Gärtner-Ellis theorem [34, 44], if $\lambda(M)$ exists and is differentiable for all M, then $P(p_\omega = \mu)$ satisfies a large deviation principle with a rate function given by

$$I(\mu) = \sup_{M \in \mathbb{R}^m} [M \cdot \mu - \lambda(M)]$$

This defines the rate function as the Legendre-Fenchel transform of λ. The relation can also be inverted: $\lambda(M) = \sup_{\mu \in \mathcal{X}} [M \cdot \mu - I(\mu)]$. We will not prove this theorem here, but heuristically, one can argue that

$$\left\langle e^{NM \cdot p_\omega} \right\rangle \asymp \int d\mu e^{N[M \cdot \mu - I(\mu)]}$$

Similar as before, the saddle-point approximation of Laplace tells us that

$$\left\langle e^{NM \cdot p_\omega} \right\rangle \asymp \exp\{N \sup_{\mu \in \mathcal{X}} [M \cdot \mu - I(\mu)]\}$$

The Gärtner-Ellis theorem can help for explicit calculations of rate functions, as it is sometimes easier to calculate the generating function λ. Vice versa, sometimes it is easier to compute a rate function to get information about a generating function.

As an illustrative example, we assume again that $\rho_k = 1/m$. Consider the following generating function:

$$\lambda(E_k) = \lim_{N\to\infty} \frac{1}{N} \log \left\langle e^{-N\beta \sum_k E_k \mu_k} \right\rangle$$

As we already know the rate function $I(\mu)$ in this case, we can compute the generating function by

$$\lambda(E_k) = \sup_{\mu} [-\beta \sum_k E_k \mu_k - I(\mu)]$$

We find that the canonical distribution $\mu_k = \frac{1}{Z} e^{-\beta E_k}$ solves this variational problem. As a consequence, the generating function $\lambda(E_k) = \log Z$. Remember that we interpret the rate function $I(\mu)$ as the entropy per particle $S/(Nk_B)$, and $-\frac{1}{\beta} \log Z$ is the free energy F. The Gärtner-Ellis theorem thus relates these two by the Legendre transform, schematically:

$$F = \sup_{\mu} [\langle E \rangle_\mu - TS(\mu)]$$

9.5 Typical behaviour and small fluctuations

The large deviation theory incorporates the law of large numbers and the central limit theorem in the following sense.

Law of large numbers Suppose that the rate function $I(\mu)$ has a unique global minimum at $\mu = \rho$. First of all, we can show that this minimum should satisfy $I(\rho) = 0$. This follows from the fact that $\lambda(0) = 0$:

$$0 = \lambda(0) = \sup_{\mu} [-I(\mu)] = -I(\rho)$$

Secondly we see that for any (measurable) set $B \subset \mathcal{X}$:

$$P(p_\omega \in B) \asymp e^{-N \inf_{\mu \in B} I(\mu)}$$

This tells us that $P(p_\omega \in B) \to 0$ whenever $\rho \notin B$, and $P(p_\omega \in B) \to 1$ whenever $\rho \in B$. This is the law of large numbers, with ρ the typical value to which p_ω converges.

Central limit theorem Given again a rate function $I(\mu)$ with a unique global minimum at ρ. Suppose that the rate function is twice differentiable at ρ. We define 'small fluctuations' as values μ close enough to ρ so that we can approximate the rate function by

$$I(\mu) \approx \frac{1}{2} \sum_{k,l} \left. \frac{\partial^2 I}{\partial \mu_k \partial \mu_l} \right|_{\mu=\rho} (\mu_k - \rho_k)(\mu_l - \rho_l)$$

In this approximation, the probability density has the form of a Gaussian for large N:

$$P(p_\omega = \mu) \approx \exp\{-\frac{N}{2} \sum_{k,l} \left. \frac{\partial^2 I}{\partial \mu_k \partial \mu_l} \right|_{\mu=\rho} (\mu_k - \rho_k)(\mu_l - \rho_l)\}$$

Which is (a form of) the central limit theorem. For this reason fluctuations in this regime are also called small fluctuations, Gaussian fluctuations or normal fluctuations.

For our specific example, where $I(\mu) = \sum_k \mu_k \log \frac{\mu_k}{\rho_k}$ we have that

$$\left. \frac{\partial^2 I}{\partial \mu_k \partial \mu_l} \right|_{\mu=\rho} = \delta_{k,l} \frac{1}{\rho_k}$$

so that for small fluctuations

$$P(p_\omega = \mu) \approx \exp\{-\frac{N}{2} \sum_k \frac{(\mu_k - \rho_k)^2}{\rho_k}\}$$

which coincides with (9.2) as expected.

9.6 Out of equilibrium

Because large deviation theory gives such a natural framework for equilibrium statistical mechanics, it is a natural candidate to give an extension to nonequilibrium. As we said in the introduction, stochastic processes are an important tool in nonequilibrium statistical mechanics.

The (mathematical) large deviation theory for Markov processes was thoroughly founded by Donsker and Varadhan in [27, 28, 29, 30]. In the context of physics, Onsager and Machlup [87] already used the large deviation approach in the 1950's

to examine relaxation to equilibrium in linear diffusion systems. A more general treatment of this problem, with a different emphasis, was done by Freidlin and Wentzell [41, 103]

In more recent years advances have been made in what is called static fluctuation theory [5, 6]. This is set in nonequilibrium stationary dynamics, and one imagines the system to be initially in the stationary regime a long time ago (at time $-\infty$, the distribution of states was the stationary distribution). The probability is then considered that the distribution at time zero is equal to some other given distribution (a fluctuation).

In this thesis we discuss dynamical fluctuation theory, in which time-integrated observables are considered, with time as the large parameter. This is mathematically not a very big step: it boils down to renaming the particle number N to the time T. For example: consider the vector $\omega = (\omega_1, \ldots, \omega_N)$ from the example above. It describes the energies of the different particles. We could also see it as the consecutive energies of one particle through time, so that ω becomes a trajectory like we have used throughout this text. The empirical occupation vector p_ω then gives us the fractions of time that the particle has spent in each energy state. The only mathematical difference is that we work with continuous time. This corresponds to measurements that are made continuously. We clarify this in the next chapters. Apart from having a different interpretation, this switch to time as a large parameter also allows us to consider small systems, without having to make some hydrodynamic approximations.

In [70, 71, 72, 96] we have reported the application of dynamical fluctuation theory on various classes of models. The major goal in this line of research is similar to that of fluctuation-dissipation relations: to find out what thermodynamic quantities govern the rate functions of the large deviations. Again, entropy and traffic turn out to be the basic ingredients for constructing this out-of-equilibrium theory.

We restrict ourselves in the next two chapters to the work reported in [70, 71] in which respectively Markov jump processes and overdamped diffusions are treated. However, the overall strategy in all models we have considered is the same: one starts by correctly defining the observables (the time integrated occupations and currents, as is explained there). Then, the rate functions (fluctuation functionals) are computed for the joint probability of occupations and currents, which can be done explicitly. The rate functions for occupations or currents alone can then be computed by the contraction principle, which is unfortunately not always possible. Throughout this calculation the physical aspect is most important: we keep track of the influence of entropy and traffic on the rate functions. We also show that the large deviation theory provides a way of defining thermodynamic potentials, although these potentials do not directly correspond to the ones defined in equilibrium thermodynamics. When only small deviations from the stationary

behaviour are considered, the rate functions are approximated to a quadratic dependence on the fluctuations (the regime where the central limit theorem applies). When the dynamics is also close to equilibrium, known variational principles such as the minimum entropy production principle are recovered.

Chapter 10

Driven overdamped diffusions

We apply the theory of large deviations to time integrated observables in general overdamped diffusion processes. These observables are empirical occupation densities and empirical currents. Central is the joint rate function for occupations and currents and the role of entropy and traffic in it. This chapter represents work reported in [71].

10.1 Model

We quickly recapitulate the equations describing overdamped diffusion, which we discussed in Section 4.5.

10.1.1 Overdamped diffusions

General overdamped diffusions in d dimensions are governed by the Langevin equation (4.12), interpreted in the Itô way:

$$dx_t^i = \sum_j [\chi_{ij}(x_t)F^j(x_t)dt + \frac{\partial D_{ij}}{\partial x^j}(x_t)dt + \sqrt{2D(x_t)}_{ij}dB_t^j] \tag{10.1}$$

where x^i denotes a spatial component of the position x. In the shorter vector/matrix notation this gives:

$$dx_t = \chi(x_t)[f(x_t) - \nabla U(x_t)]dt + \nabla \cdot D(x_t)dt + \sqrt{2D(x_t)}dB_t \tag{10.2}$$

In this equation we have explicitly written the energy U of the system, giving a conservative forcing, while f_t represents a nonconservative force (i.e. the driving). Furthermore D and χ are $d \times d$ symmetric matrices both depending on the position x_t, and there are d Wiener processes dB_t^j. We restrict ourselves to the case that the system is in contact with one heat bath at a temperature β. Remember that the local detailed balance condition implies $\chi = \beta D$.

We restrict ourselves to two types of boundary conditions:
(1) *periodic*—the particle moves on the unit torus $[0,1)^d$ and the fields U, f, and χ are smooth functions on the torus;
(2) *decay at infinity*—the potential U grows fast enough at infinity so that the particle is essentially confined to a bounded region, i.e., the density and its derivative vanish at infinity.

Under either of the above boundary conditions we can simply ignore boundary terms when performing integrations by parts. The particles are essentially confined in their configuration space.
The probability density μ_t evolves according to the Fokker-Planck equation (4.14):

$$\frac{\partial \mu_t}{\partial t} + \nabla \cdot j_{\mu_t} = 0, \qquad j_\mu = \chi\mu\,(F - \nabla U) - D\nabla\mu \tag{10.3}$$

The stationary condition reads $\nabla \cdot j_\rho = 0$.

This class of models is still quite large. It is illustrative to consider a simple example of these dynamics: a particle moving on a circle (one dimension). The advantage here is that we can explicitly calculate the stationary distribution.

10.1.2 Diffusion on a circle

We consider a particle undergoing an overdamped motion on the circle with unit length:

$$dx_t = \chi(x_t)\big[f(x_t) - U'(x_t)\big]\,dt + D'(x_t)dt + \sqrt{2D(x_t)}\,dB_t \tag{10.4}$$

The prime as superscript is a shorthand notation for the spatial derivative.

The corresponding Fokker-Planck equation for the time-dependent probability density μ_t is

$$\frac{\partial \mu_t(x)}{\partial t} + j'_{\mu_t}(x) = 0, \qquad j_\mu = \chi\mu(f - U') - D\mu' \tag{10.5}$$

The stationary density ρ solves the stationary equation $j'_\rho = 0$, i.e.,

$$\chi\rho(f - U') - D\rho' = j_\rho \tag{10.6}$$

is a constant. For $f = 0$, equation (10.6) has the solution

$$\rho(x) = \frac{1}{Z}\, e^{-\beta U(x)}, \qquad Z = \int_0^1 e^{-\beta U} dx \tag{10.7}$$

and the corresponding stationary current is $j_\rho = 0$; this is a detailed balanced dynamics with ρ the equilibrium density.
When adding a nongradient driving force, $\int_0^1 f dx \neq 0$, (10.6) can still be solved; the stationary density obtains the form

$$\rho(x) = \frac{1}{\mathcal{Z}} \int_0^1 \frac{e^{\beta W(y,x)}}{D(y)}\, dy, \qquad \mathcal{Z} = \int_0^1 \int_0^1 \frac{e^{\beta W(y,x)}}{D(y)}\, dydx \tag{10.8}$$

where

$$W(y,x) = U(y) - U(x) + \begin{cases} \int_y^x f\, dz & \text{for } y \leq x \\ \int_y^1 f\, dz + \int_0^x f dz & \text{for } y > x \end{cases} \tag{10.9}$$

is the work performed by the applied forces along the positively oriented path $y \to x$. In this model the stationary current can be computed by dividing the stationary equation (10.6) by $\rho\chi$ and by integration over the circle:

$$j_\rho = \frac{\overline{W}}{\int_0^1 (\rho\chi)^{-1} dx} \tag{10.10}$$

where $\overline{W} = \int_0^1 f dx$ is the work carried over a completed cycle. The non-zero value of this stationary current indicates that time-reversibility is broken. In the simplest nonequilibrium setting when $U = 0$ and $f, \chi > 0$ are some constants, the steady state has the uniform density $\rho(x) = 1$ and the current is $j_\rho = \chi f$.

10.2 Entropy and traffic

Before going on to defining the observables for which we want to describe large deviation principles, let us first discuss the basic ingredients that play a role in the rate functions: entropy and traffic.

To properly define them, we need to define a reference process. We take this process to be (see 4.7):

$$dx_t = \nabla \cdot D(x_t)dt + \sqrt{2D(x_t)}dB_t$$

i.e. the overdamped diffusion (10.2) with the force $f - \nabla U$ put to zero. The action that describes the probabilities of trajectories $\omega = (x_t)_{0\leq t\leq T}$ of the original process with respect to the reference process is given by (4.15):

$$\begin{aligned} e^{-A(\omega)} &= \frac{d\mathcal{P}_{\mu_0}}{d\mathcal{P}^0_{\mu_0}}(\omega) \qquad (10.11)\\ &= \exp\left\{\frac{1}{4}\int_0^T dt[2\dot{x}_t + 2\nabla\cdot D - \chi(f-\nabla U)]\cdot D^{-1}\chi(f-\nabla U)\right\} \end{aligned}$$

10.2.1 Entropy

The entropy flux into the environment is equal to the time-antisymmetric part of the action (by local detailed balance, see (4.22)):

$$\begin{aligned} S(\omega) = A(\theta\omega) - A(\omega) &= \beta\int dx_t \circ (f(x_t) - \nabla U(x_t))\\ &= \beta\int dx_t \circ f(x_t) - \beta[U(x_T) - U(x_0)] \qquad (10.12) \end{aligned}$$

which is equal to the work done by the nonconservative force (see (4.21)) minus the change of energy of the system. It is useful for later calculations to compute the average of the entropy flux, started from some initial distribution μ_0. For that we need to compute the average of the Stratonovitch integral. We rewrite the Stratonovitch integral to an Itô integral using (4.19):

$$\int_0^T dx_t \circ f(x_t) = \int_0^T dx_t f(x_t) + \int_0^T dt[D(x_t)\nabla]\cdot f(x_t)$$

Using the Langevin equation, the average of the Itô integral $\left\langle \int_0^T dx_t f(x_t)\right\rangle_{\mu_0}$ gives:

$$\begin{aligned} &\int_0^T \Big\langle f(x_t)\cdot\Big(\chi(x_t)[f(x_t)-\nabla U(x_t)]dt + \nabla\cdot D(x_t)dt + \sqrt{2D(x_t)}dB_t\Big)\Big\rangle_{\mu_0}\\ &= \int_0^T dt\Big\langle f(x_t)\Big(\chi(x_t)[f(x_t)-\nabla U(x_t)] + \nabla\cdot D(x_t)\Big)\Big\rangle_{\mu_0} \end{aligned}$$

In the last step we used that the average over the term with dB_t drops, because $dB_t = B_{t+dt} - B_t$ is independent of x_t and has average zero. Then, by definition of the time-evolved distribution μ_t we get

$$\left\langle \int_0^T dx_t \circ f(x_t) \right\rangle_{\mu_0} = \int_0^T dt \int dx \mu_t \Big(f \cdot \chi[f - \nabla U] + \nabla \cdot [Df] \Big)$$

which can be rewritten succinctly, using partial integration, as

$$\left\langle \int_0^T dx_t \circ f(x_t) \right\rangle_{\mu_0} = \int_0^T \int dx dt f \cdot j_{\mu_t} \tag{10.13}$$

where j_{μ_t} is the probability current as in (10.3). The average entropy flux, starting from a density μ_0 is thus equal to

$$\langle S(\omega) \rangle_{\mu_0} = \int_0^T dt \left[\int dx f \cdot j_{\mu_t} - \frac{d}{dt} \int dx U(x) \mu_t(x) \right] \tag{10.14}$$

In Section 2.8 we also discussed the average entropy. There we defined the entropy of the system through the Shannon (or Gibbs) entropy:

$$s(\mu) = - \int dx \mu(x) \log \mu(x)$$

The total average change of entropy (entropy production) in the world, given the initial distribution μ_0 is then, see (2.18):

$$\langle S_{\mu_0}(\omega) \rangle_{\mu_0} = \int_0^T dt \sigma(\mu_t) \tag{10.15}$$

where the instantaneous entropy production rate $\sigma(\mu)$ is easily computed to be

$$\sigma(\mu) = \int j_\mu \cdot (\mu D)^{-1} j_\mu \, dx \tag{10.16}$$

It is very important to note that, although computed through the action relative to a reference process, the average entropies computed here do not depend on that reference process (as long as the reference process is an equilibrium process).

10.2.2 Traffic

The traffic is defined as the time-symmetric part of the action (10.11):

$$\begin{aligned} \mathcal{T}(\omega) &= A(\theta\omega) + A(\omega) \qquad (10.17) \\ &= \frac{\beta^2}{2}\int_0^T dt(f-\nabla U)\cdot D(f-\nabla U) + \beta\int_0^T dt\nabla\cdot[D(f-\nabla U)] \end{aligned}$$

The average value of the traffic, given an initial distribution μ_0 is then equal to

$$\langle \mathcal{T}(\omega)\rangle_{\mu_0} = \int_0^T dt\tau(\mu_t)$$

with the instantaneous 'traffic rate' given by

$$\tau(\mu) = \frac{\beta^2}{2}\int dx\mu(f-\nabla U)\cdot D(f-\nabla U) + \beta\int dx\mu\nabla\cdot[D(f-\nabla U)] \qquad (10.18)$$

The traffic, in contrast to entropy, is very much dependent on the choice of the reference process. It is interesting to note that for overdamped diffusions, the entropy rate and traffic rate are closely related. To see this, note that the entropy production rate for the reference process, can be found by simply putting $f-\nabla U$ to zero in the explicit formula (10.16):

$$\sigma^0(\mu) = \int dx\frac{(\nabla\mu)\cdot D(\nabla\mu)}{\mu}$$

One can then check directly that

$$\tau(\mu) = \frac{\sigma(\mu)-\sigma^0(\mu)}{2} \qquad (10.19)$$

In following sections we often compare the traffic rates of different dynamics (determined by different forces). This equality then tells us that these traffic differences can be replaced by entropy differences. Hence, relation (10.19) brings about a simplification in the structure of fluctuations that is characteristic and restricted to diffusions. See for example the next chapter, where there is no relation like (10.19). We believe that this also indicates that a more general nonequilibrium theory should reach beyond the Langevin or diffusion approximation.

10.3 Observables

Here we define the observables for which we want to compute the rate functions.

A basic and time-symmetric dynamical observable is the empirical distribution of the occupation times. It is defined as the fraction of time spent in a subset A of the state space Ω over a fixed time interval T:

$$p_\omega(A) = \frac{1}{T}\int_0^T \chi_A(x_t)\,dt$$

where $\chi_A(x)$ is the indicator function giving a value 1 if $x \in A$ and zero otherwise. Note that this resembles the definition of the empirical occupation vector (9.1) in the example explained in the last chapter. The difference is that the particle number N is replaced by the duration time T and the state space is continuous.

Having this in mind we formally write the empirical occupation density p:

$$p_\omega(x) = \frac{1}{T}\int_0^T \delta(x_t - x)\,dt \tag{10.20}$$

This is a path-dependent observable as it varies over the paths $\omega = (x_t)_{0\le t\le T}$. Note that it is normalized (i.e. integration over state space gives one). It is also a useful observable, because we can write time-integrated state-functions g in terms of this empirical density:

$$\frac{1}{T}\int_0^T dtg(x_t)\,dt = \int dxp_\omega(x)g(x)$$

Because we always assume that there exists a unique stationary distribution, we have ergodicity:

$$\int dxp_\omega(x)g(x) \to \langle g\rangle_\rho$$

almost surely for $T \to \infty$ for any state function g (this means that the probability to have a path ω for which this convergence holds is equal to one). We conclude that for large times $p_\omega \to \rho$ almost surely.

The time-antisymmetric observable of special relevance is the empirical current. We define the total time-averaged current in a set $A \subset \Omega$ as

$$J_\omega(A) = \frac{1}{T}\int_0^T dx_t \circ \chi_A(x_t)$$

where the circle denotes a Stratonovitch stochastic integral, and χ_A is again an indicator function. This observable is the sum of all displacements the particle makes when in a subset A of Ω and thus represents the time-integrated particle current (per unit of time). As usual, there can be an ambiguity in choosing the type of stochastic integral. The Stratonovitch integral is there to assure that the observable is time-antisymmetric. The current density is then formally defined as

$$J_\omega(x) = \frac{1}{T}\int_0^T dx_t \circ \delta(x_t - x) \tag{10.21}$$

It depends again on the (random) path ω and it measures the time-averaged current while in x. As for the empirical occupations, this observable is useful because quantities like work and the entropy flux can be written in terms of it: for an arbitrary state function g we get

$$\frac{1}{T}\int_0^T dx_t \circ g(x_t) = \int dx J_\omega(x) g(x)$$

In the same way as for occupations and with (10.13) we see that $J_\omega \to j_\rho$ for $T \to \infty$ almost surely.

10.4 Fluctuations of occupations and currents

We already argued that the empirical densities of occupations and currents will, in the large time limit, converge almost surely to the stationary densities ρ and j_ρ. In this section we compute the rate functions (fluctuation functionals) for deviations from these typical values. This amounts to computing the probabilities that the empirical observables p_ω, J_ω are close to some given but arbitrary density μ and current j.

10.4.1 Definitions and restrictions

To be able to compute the probabilities of the empirical observables, we have to define the distance d between p_ω and μ and between J_ω and j. With such a definition, probabilities like $P(d(p_\omega, \mu) < \epsilon)$ for some small number $\epsilon > 0$ should be computed. Actually proving that the large deviation principle holds, i.e.

$$-\lim_{T\to\infty} \frac{1}{T} \log P(d(p_\omega, \mu) < \epsilon) = I(\mu) + O(\epsilon) \tag{10.22}$$

and under what restrictions, is not trivial. In this text we do not concern ourselves with this. For mathematical details, see [24, 27, 28, 29, 30]. Instead, we assume that the large deviation principle holds, and concentrate on computing the rate functions using heuristic arguments. We begin by introducing the following simplified notation for (10.22):

$$-\lim_{T\to\infty}\frac{1}{T}\log P(p_\omega=\mu)=I(\mu)\qquad\text{or}\qquad P(p_\omega=\mu)\asymp e^{-TI(\mu)}$$

Similarly as for the occupations the probability of deviations for the currents are defined. However, the natural starting point is the joint probability for currents and occupations:

$$P(p_\omega=\mu,J_\omega=j)\asymp e^{-TI(\mu,j)}$$

as both $I(\mu)$ and $I(j)$ can be obtained in principle from the joint fluctuations by the contraction principle.

A first observation that we can make, is that $I(\rho,j_\rho)=0$, i.e. the stationary density and current have a probability that does not decay to zero. Moreover, the stationary regime is characterized by

$$I(\rho,j_\rho)=\inf_{\mu,j}I(\mu,j)$$

A second observation is that $I(\mu,j)=\infty=I(j)$ whenever j is not stationary, i.e. for $\nabla\cdot j\neq 0$. Indeed, for any smooth bounded function Y one has

$$\int Y\,\nabla\cdot J_\omega\,dx=-\frac{1}{T}\int_0^T\nabla Y(x_t)\circ dx_t=-\frac{1}{T}[Y(x_T)-Y(x_0)]\to 0$$

for $T\to\infty$. Hence, in a distributional sense, $\nabla\cdot J_\omega\to 0$ for $T\to\infty$ along *any* particle trajectory, which proves the above statement. That is why from now on we always assume that $\nabla\cdot j=0$, unless otherwise specified.

10.4.2 Computation of the fluctuation functional

The probability of the fluctuations can be written as

$$P(p_\omega=\mu,J_\omega=j)=\int dP(\omega)\chi[p_\omega=\mu,J_\omega=j]$$

where χ is again an indicator function, giving one when its argument is satisfied and zero otherwise. We have omitted initial conditions in the notation, as they

are not important in the long-time limit. (To be more rigorous, one has to define a neighbourhood of μ and j using the distance defined on the function space. The indicator function then gives one when the empirical densities are in that neighbourhood.) To compute the rate function we exploit the knowledge we have of Radon-Nikodym derivatives. For this we define a new dynamics, by changing the force f in (10.2) to a new force g. We take this g such that the density μ and the current j become typical, i.e:

$$j = \chi[g - \nabla U]\mu - D\nabla\mu$$

This explicitly defines g. With this new dynamics we write:

$$P(p_\omega = \mu, J_\omega = j) = \int dP_g(\omega) \frac{dP_f}{dP_g}(\omega) \chi[p_\omega = \mu, J_\omega = j] \tag{10.23}$$

where we now have added subscripts f and g to denote respectively the old and the new dynamics. The Radon-Nikodym derivative can be written in terms of the entropy production and traffic:

$$\log \frac{dP_f}{dP_g}(\omega) = S_f(\omega) - S_g(\omega) + \mathcal{T}_g(\omega) - \mathcal{T}_f(\omega)$$

which are defined in (10.12) and (10.17). Again the quantities with a subscript g are defined in the dynamics with f replaced by g. The excess entropy can be rewritten in terms of the empirical current:

$$S_f(\omega) - S_g(\omega) = \beta \int dx_t \circ [f(x_t) - g(x_t)] = \beta T \int dx J_\omega(x)[f(x) - g(x)]$$

The excess traffic can be rewritten in terms of the empirical occupations:

$$\mathcal{T}_g(\omega) - \mathcal{T}_f(\omega) = \beta T[\tau_g(p_\omega) - \tau_f(p_\omega)]$$

where τ_g is given by (10.18), but with f replaced by g. We see that the entropy production in both dynamics only depends on the path ω through the empirical current, and the traffic in both dynamics only depends on the empirical occupation density. When we substitute this in (10.23) we see that the indicator function allows us to replace the empirical observables by the density μ and the current j. We thus get

$$P(p_\omega = \mu, J_\omega = j) = e^{-TI(\mu,j)} \int dP_g(\omega) \chi[p_\omega = \mu, J_\omega = j] \tag{10.24}$$

where the functional $I(\mu, j)$ is given by

$$I(\mu, j) = \frac{1}{2}\tau_f(\mu) - \frac{1}{2}\tau_g(\mu) + \frac{\beta}{2}\int dxj \cdot (g - f) \tag{10.25}$$

This functional is exactly the rate function, because the last factor in (10.24) goes to one in the limit of large times, which is a consequence of the fact that μ and j are the typical (stationary) values in the dynamics determined by g.

Let us examine this joint rate function $I(\mu, j)$. It is the sum of an excess instantaneous traffic rate given the density μ alone, and an excess work (or, equivalently, entropy flux) given stationary current j alone. The excess is an excess of the original dynamics with respect to the dynamics in which μ and j are typical. If we use (10.19), we can rewrite the functional in terms of the entropy production rate:

$$I(\mu, j) = \frac{1}{4}\sigma_f(\mu) - \frac{1}{4}\sigma_g(\mu) + \frac{\beta}{2}\int dxj \cdot (g - f) \tag{10.26}$$

Furthermore, using the explicit expressions of the traffic/entropy and the force g we can rewrite the fluctuation functional in the following explicitly positive form:

$$\begin{aligned} I(\mu, j) &= \frac{1}{4}\sigma_f(\mu) - \frac{1}{4}\sigma_g(\mu) + \frac{\beta}{2}\int dxj \cdot (\mu D)^{-1}(j - j_\mu) \\ &= \frac{1}{4}\int (j - j_\mu) \cdot (\mu D)^{-1}(j - j_\mu)\, dx \end{aligned} \tag{10.27}$$

(On the assumption $\nabla \cdot j = 0$; remember that $I(\mu, j) = \infty$ otherwise.) This last formula resembles the Gaussian-like expressions for the current distribution, typical for hydrodynamic fluctuations of the diffusion-type. Such expressions are omnipresent in the works of e.g. [7, 9]. Although the quadratic integrand in (10.27) resembles the (generalized Onsager-Machlup) Lagrangian for macroscopic fluctuations in the hydrodynamic limit, we have no spatial/temporal rescaling here. We have started from a mesoscopic system as described by a diffusion equation and the only large parameter is the time span T. The difference between our work and other work on dynamical large deviations becomes visible from the formula (10.25) and (10.27). Our approach has nothing to do with hydrodynamic rescaling or with macrostatistics. It concerns the thermodynamic interpretation of the fluctuation functional $I(\mu, j)$ for our mesoscopic system: how it is shaped from quantities like traffic, work and entropy production, and providing full account of the steady dynamical fluctuations in both the time-symmetric and the time-antisymmetric sectors.

10.5 A notion of thermodynamic potentials

In equilibrium systems it is useful to consider thermodynamic potentials (like free energy), as they have a clear physical meaning and characterize equilibrium via variational principles. Moreover, to go from one potential to another, Legendre transforms are used. Fluctuation functionals also bring with them variational principles. First of all, and mainly, minimizing them characterizes stationarity. In (10.25) we gave an explicit expression for the rate function of occupations and currents. Let us analyze this from the viewpoint of thermodynamic potentials and Legendre transforms. First of all, the first two terms in (10.25) constitute an excess traffic, or equivalently excess entropy production (10.19). Let us consider the traffic in a dynamics governed by an arbitrary force h:

$$\tau_h(\mu) = \frac{\beta^2}{2} \int dx \mu (h - \nabla U) \cdot D(h - \nabla U) + \beta \int dx \mu \nabla \cdot [D(h - \nabla U)]$$

This can be seen as a potential for the currents in the sense that the functional derivative of it with respect to the force, gives

$$\frac{\delta \tau_h}{\delta h(x)} = \beta j^h_\mu(x) \tag{10.28}$$

where j^h_μ is the probability current in a dynamics with a force h, see (10.3). It is then natural to examine the Legendre transform of τ_h:

$$G(\mu, j) = \sup_h \left[\beta \int h \cdot j dx - \tau_h(\mu) \right]$$

We find the supremum by using Euler-Lagrange equations:

$$j = j^h_\mu$$

so that $h = g$ is exactly the force needed to make the current j typical, together with μ:

$$G(\mu, j) = \beta \int g \cdot j dx - \tau_g(\mu)$$

On the other hand, if we take the functional derivative of G, we get:

$$\frac{\delta G}{\delta j(x)} = \beta g(x) \tag{10.29}$$

where g is again the force that makes μ and j typical. So G is a potential for the forces, just like τ was a potential for the currents, and by Legendre transforms we can switch between the two.

We can easily rewrite $I(\mu, j)$ in terms of G and τ:

$$I(\mu, j) = \frac{1}{2}\left[G(\mu, j) + \tau_f(\mu) - \beta \int dx f \cdot j\right] \tag{10.30}$$

Note that $\tau_0(\mu) = 0$, so that $\frac{1}{2}G(\mu, j)$ is exactly equal to the fluctuation functional in the case that $f = 0$. In other words:

$$I_f(\mu, j) = I_0(\mu, j) + \frac{1}{2}\left[\tau_f(\mu) - \int dx f \cdot j\right]$$

This is nice, because the left-hand side is a fluctuation functional for a nonequilibrium dynamics, while on the right-hand side, the first term is a fluctuation functional in an equilibrium dynamics. The rest is thus 'the correction to equilibrium.'

We can still rewrite (10.30) in the form

$$4I_f(\mu, j) = \sup_f \{2\beta \int dx f \cdot j - \sigma_f(\mu)\} - 2\beta \int dx f \cdot j + \sigma_f(\mu) \tag{10.31}$$

fully in terms of entropic quantities, due to (10.19). A similar structure, cf. (10.28) and (10.29), has been established already before in the framework of jump processes, see [69] and also the next chapter.

10.6 Contractions

Now that we have a fluctuation functional for both symmetric and antisymmetric variables, we can compute the statistics of empirical averages of arbitrary physical quantities. In particular, we can try to find the fluctuation functionals for density $I(\mu)$ and for current $I(j)$ separately. To start we look at the fluctuations of the occupation density.

10.6.1 Occupation statistics

As $I(\mu) = \inf_j I(\mu, j)$, we have to compute the minimizing current j for any given density μ. Since the minimization is constrained via the stationary condition $\nabla \cdot j = 0$, we get the Euler-Lagrange equation from (10.27)

$$j = \chi\mu\big[f - \nabla \cdot (U + V)\big] - D\nabla\mu \tag{10.32}$$

where V is a Lagrange multiplier (function of x). Not surprisingly, we see that the minimizer is the stationary current for a modified dynamics that makes μ stationary. This modified dynamics is achieved here by adding to the imposed potential U an extra potential V. We therefore call the minimizing current in (10.32) j^V_μ, and the fluctuation functional becomes:

$$I(\mu) = \frac{1}{4}\int (j^V_\mu - j_\mu) \cdot (\mu D)^{-1}(j^V_\mu - j_\mu)\, dx \tag{10.33}$$

For some explicit examples of solutions to (10.32), see further down, in equations (10.37) and (10.46).

The fluctuation functional $I(\mu)$ obtains other equivalent forms by substituting $g = f - \nabla V$ into (10.25):

$$I(\mu) = \frac{\tau(\mu) - \tau_V(\mu)}{2} = \frac{\sigma(\mu) - \sigma_V(\mu)}{4} \tag{10.34}$$

where the second equality follows again from (10.19). In this way we have recognized the excess traffic (or here also: the excess entropy production) as governing the large time statistics of the occupation times. Excess is here excess of the original with respect to the modified process in which μ is typical. The fluctuation functional $I(\mu)$ thus exactly equals one quarter of a difference in entropy production rates when having density μ, these rates being computed respectively for the original dynamics and for a modified dynamics that makes μ stationary.

In formulæ (10.33) and (10.34) the potential V has to be determined from μ by solving the inverse stationary problem (10.32). We discuss two examples in which this can be done explicitly: the equilibrium case and the example of diffusion on the circle.

Equilibrium dynamics Let us see what can be said in general for equilibrium diffusions. If $f = 0$ then equation (10.32) has the solution $V = -\beta^{-1}\log\mu$, and

the corresponding current j^V_μ and the entropy production $\sigma_V(\mu)$ are both zero. As a result,

$$I(\mu) = \frac{\sigma(\mu)}{4} \tag{10.35}$$

This exact relation between the equilibrium dynamical fluctuations and the entropy production is solely true for diffusion processes. In contrast, for jump processes (see the next chapter) $\sigma(\mu)$ gives only the leading term in an expansion of $I(\mu)$ around the equilibrium density $\rho \propto e^{-\beta U}$, and the relation (10.35) obtains corrections when beyond small fluctuations; see [68] for details.

Diffusion on the circle For the one-dimensional example of Section 10.1.2 the inverse stationary problem (10.32) allows for an explicit solution. The current j^V_μ is immediately read off the formula (10.10),

$$j^V_\mu = \frac{\beta\overline{W}}{\int_0^1 (\mu D)^{-1} dx}, \qquad \overline{W} = \int_0^1 f\, dx \tag{10.36}$$

and the potential V obtains the form

$$V(x) = -\frac{1}{\beta}\log\mu(x) + \int_0^x \big(f - \frac{j^V_\mu}{\beta\mu D}\big)\, dy \tag{10.37}$$

which is a nonlocal functional of the given density μ. The fluctuation functional is explicitly given as

$$4I(\mu) = \sigma(\mu) - \frac{\overline{W}^2}{\int_0^1 (\mu D)^{-1} dx} \tag{10.38}$$

for $\mu \neq 0$.
Observe that if $\mu = 0$ on some open set A then the rate function equals $I(\mu) = \sigma(\mu)/4$. (That follows also from the equilibrium form (10.35) below as the circle gets *effectively* cut and the dynamics mimics a detailed balance one.) The infimum of $I(\mu)$ over all densities μ that vanish on A then gives the escape rate from the complement $A^c = [0,1) \setminus A$.
As a simple example, assume that $U = 0$ and let f and D be some constants. In this case the entropy production (10.16) reads

$$\sigma(\mu) = \beta^2 D f^2 + D\int_0^1 \frac{\mu'^2}{\mu}\, dx \tag{10.39}$$

To compute the escape rate from A^c (or, entrance rate to A) we must take the infimum of (10.39) over all μ that vanish on A. Setting $A = (0, \delta)$ for some $0 < \delta < 1$, that infimum is reached for the density $\mu^*(x) = \frac{2}{1-\delta} \sin^2(\frac{\pi(x-\delta)}{1-\delta}), x \in [\delta, 1]$, and the escape rate is

$$\inf_{\mu|_A=0} I(\mu) = I(\mu^*) = \frac{\pi^2 D}{(1-\delta)^2} + \frac{\beta^2 D f^2}{4} \tag{10.40}$$

Even in equilibrium ($f = 0$) the result is meaningful as it relates the diffusion constant to an escape rate. In the context of dynamical systems, the analysis of the escape rates and of their link to linear transport coefficients was initiated by Dorfman and Gaspard, see [31, 32] and references therein.

10.6.2 Current statistics

The contraction to the current j is also possible. However, up to special examples, there is no explicit solution to the associated variational problem and for general models one has to resort to a perturbative or numerical analysis. In fact, often the calculation starting from the generating function of the current appears more practical than to do the contraction starting from $I(\mu, j)$, see e.g. [25]. Through the Gärtner-Ellis theorem the generating function is related to the fluctuation functional via a Legendre transform.

However, we can check in general a fluctuation theorem for currents: notice that

$$I(\mu, -j) - I(\mu, j) = \frac{1}{2} \int dx j (\mu D)^{-1} j_\mu = \frac{\beta}{2} \int dx j \cdot f$$

This remains true even when the contraction to currents alone is made. We thus have that

$$\frac{P(J_\omega = j)}{P(J_\omega = -j)} \asymp \exp\left\{ \frac{\beta T}{2} \int dx j \cdot f \right\}$$

We see that currents that have on average the same direction as the forcing, are exponentially more probable than their reversed currents.

For explicit computations of the current rate function we restrict us here to giving the result for a constant drift on the circle.

Constantly driven diffusion on the circle We take $U = 0$ and f, χ constants. In this case, from (10.27) the joined fluctuation functional reads:

$$I(\mu, j) = \frac{1}{4D} \int \frac{1}{\mu}(j - \beta D f \mu - D\mu')^2 \, dx \tag{10.41}$$

and for all j, the infimum over μ is reached at the uniform distribution, so that

$$I(j) = \frac{(j - \beta D f)^2}{4D} \tag{10.42}$$

and hence we see that here the current fluctuations are Gaussian.

10.7 Small fluctuations and entropy principles

Considering small fluctuations (for which a quadratic approximation is valid) can simplify the fluctuation functionals mathematically. It is also experimentally a more accessible regime, as the fluctuations are more probable than large fluctuations. The fluctuation functionals can be expanded in both the occupation densities and currents around their typical values and the strictly positive quadratic form obtained in the leading order describes normal (Gaussian) fluctuations. From a physical point of view, the structure of these normal fluctuations have been first analyzed by Onsager and Machlup, [87], for the case of relaxation to equilibrium. Here we show a natural extension of the original Onsager-Machlup formalism to nonequilibrium systems by starting from the above fluctuation theory.

We look here at the Gaussian approximation in a dynamics far from equilibrium. Later we will also make the driving f small, to be close to equilibrium.
As is clear from (10.27), current and occupations are coupled. It is because of this coupling that contractions of $I(\mu, j)$ to $I(\mu)$ and to $I(j)$ become rather complicated. Even for small fluctuations this coupling remains: take $\mu = \rho(1 + \epsilon\mu_1)$ and $j = j_\rho + \epsilon j_1$, with ϵ a small parameter. Because $j - j_\mu$ is then $O(\epsilon)$, the fluctuation functional is $O(\epsilon^2)$:

$$\begin{aligned} I(\mu, j) &= \frac{\epsilon^2}{4} \int dx \big[j_1 \cdot (\rho D)^{-1} j_1 + \mu_1^2 j_\rho \cdot (\rho D)^{-1} j_\rho \\ &\qquad + \nabla\mu_1 \cdot \rho D \nabla\mu_1 - 2\mu_1 j_1 \cdot (\rho D)^{-1} j_\rho \big] + o(\epsilon^2) \end{aligned} \tag{10.43}$$

The last term in the integrand gives the coupling between occupation and current fluctuations. It is proportional to the stationary current, which is non-zero away

from equilibrium. It is only when we take a dynamics close to equilibrium, i.e. $f = \epsilon f_1$, that the fluctuations decouple. In this approximation we have that $j_\rho = O(\epsilon)$, and thus, near equilibrium,

$$I(\mu, j) = \frac{\epsilon^2}{4} \int dx \big[j_1 \cdot (\rho D)^{-1} j_1 + \nabla \mu_1 \cdot \rho D \nabla \mu_1 \big] + o(\epsilon^2) \tag{10.44}$$

with, to leading order, a complete decoupling between the time-symmetric (occupations) and the time-antisymmetric (current) sectors.

Occupations When close to equilibrium, the computation of $I(\mu)$ by contraction is easy: we see from (10.44) that the second term on its right-hand side is just $I(\mu)$.
In the same approximation of small fluctuations and close-to-equilibrium, the entropy production becomes:

$$\sigma(\mu) = \int dx \big[j_\rho \cdot (\rho D)^{-1} j_\rho + \epsilon^2 \nabla \mu_1 \cdot \rho D \nabla \mu_1 \big] + o(\epsilon^2) \tag{10.45}$$

and thus we get

$$I(\mu) = \frac{\sigma(\mu) - \sigma(\rho)}{4} + o(\epsilon^2) \tag{10.46}$$

This reveals to be a special case of a general result, [68], according to which the entropy production governs the occupational statistics in the linear irreversible regime. It provides a fluctuation-based explanation for the minimum entropy production principle introduced by Prigogine to characterize stationarity via an (approximate) variational principle, [89]: the stationary state has a minimal entropy production, which is not necessarily zero because the system is out of equilibrium.

Currents For currents we have an analogue of the minimum entropy production principle. The starting point is again (10.44) from which we extract the current fluctuations:

$$\begin{aligned} I(j) &= \frac{1}{4} \int (j - j_\rho) \cdot (\rho D)^{-1} (j - j_\rho) \, dx + o(\epsilon^2) \\ &= \frac{1}{4} \big[\mathcal{D}(j_\rho) + \mathcal{D}(j) - 2\mathcal{S}(j) \big] + o(\epsilon^2) \end{aligned} \tag{10.47}$$

with $\mathcal{D}(j) = \int j \cdot (\rho D)^{-1} j \, dx$ sometimes called the Onsager dissipation function, and $\mathcal{S}(j) = \beta \int f \cdot j dx$ is the entropy flux given a current j. In particular, this leads to a variational characterization of the steady current j_ρ which can be written as the following maximum entropy production principle: the j_ρ maximizes the entropy flux $\mathcal{S}(j)$ under the two stationary constraints

$$(1)\ \nabla \cdot j = 0, \qquad (2)\ \mathcal{D}(j) = \mathcal{S}(j) \tag{10.48}$$

The second condition is indeed satisfied at $j = j_\rho$ (note also that ρ can with no harm in this order be replaced by the equilibrium density $\rho_{f=0} = e^{-\beta U}/Z$.) Such a variational principle, known as a maximum entropy production principle, is often used in applications and apparently even beyond the linear irreversible regime. As is however clear from (10.47) from our dynamical fluctuation theory, the validity of the maximum entropy principle is restricted to close-to-equilibrium. Beyond that regime, we must refer to contractions from (10.27), (10.30) or even from (10.43) for generally valid expressions with a general thermodynamic meaning.

Chapter 11

Markov jump processes

Here, the large deviation theory is applied to Markov jump processes, again with empirical occupation and current densities as observables. The lines of reasoning are the same as in the last chapter, and therefore the discussion in this chapter is less detailed. For more details, we refer to [70].

11.1 Entropy and traffic

In this chapter we consider time-homogeneous Markov jump processes on a finite state space Ω, determined by the transition rates $k(x, y)$. We imagine the system to be in contact with a single heat bath. The local detailed balance assumption then restricts these rates (3.10):

$$\frac{k(x,y)}{k(y,x)} = e^{-\beta[U(y)-U(x)-W(x,y)]}$$

where $U(x)$ is the energy of the system in state x and $W(x, y) = -W(y, x)$ is the work done on the system by a nonconservative force during the transition $x \to y$.

Again the main ingredients for the dynamical fluctuation theory are the entropy and traffic, which are defined through the path-probability density, relative with respect to a reference process. As the reference process we take an equilibrium process with rates $k_0(x, y)$ defined through

$$k(x,y) = k_0(x,y)e^{\frac{\beta}{2}W(x,y)} \tag{11.1}$$

One can check that the rates $k_0(x,y)$ are detailed balanced with equilibrium distribution $\rho(x) \propto \exp[-\beta U(x)]$. The action of the original process with respect to this reference process is given by (3.9):

$$\begin{aligned} -A(\omega) &= \log \frac{d\mathcal{P}}{d\mathcal{P}_0}(\omega) \\ &= \int_0^T dt[\lambda_0(x_t) - \lambda(x_t)] + \sum_{t \le T} \log\left(\frac{k(x_{t^-}, x_t)}{k_0(x_{t^-}, x_t)}\right) \end{aligned}$$

where the sum is over all jump times of the path, x_{t^-} is the state just before the jump and $\lambda(x) = \sum_y k(x,y)$. Using the relation between the original and the reference process, this becomes

$$-A(\omega) = \int_0^T dt \sum_y k_0(x_t, y)[1 - e^{\frac{\beta}{2}W(x_t,y)}] + \frac{\beta}{2}\sum_{t \le T} W(x_{t^-}, x_t) \tag{11.2}$$

11.1.1 Entropy

The excess entropy flux (original system with respect to the reference) into the environment during a trajectory $\omega = (x_t)_{0 \le t \le T}$ is given by

$$S_{ex}(\omega) = A(\theta\omega) - A(\omega) = \beta \sum_{t \le T} W(x_{t^-}, x_t)$$

For later purposes we would like to compute the average of the entropy production. We prove that for any function of two states $g(x,y)$ and for any initial condition μ:

$$\left\langle \sum_{t \le T} g(x_{t^-}, x_t) \right\rangle_{\mu_0} = \int_0^T dt \sum_{x,y} \mu_t(x) k(x,y) g(x,y) \tag{11.3}$$

We do this by defining yet another set of rates $k^*(x,y) = k(x,y)\exp[h \cdot g(x,y)]$, with $h \in \mathbb{R}$ and using the identity $1 = \left\langle \frac{d\mathcal{P}^*}{d\mathcal{P}}(\omega) \right\rangle_{\mu_0}$:

$$1 = \left\langle \exp\left\{ \int_0^T dt \sum_y k(x_t, y)[1 - e^{hg(x_t,y)}] + h \sum_{t \le T} g(x_{t^-}, x_t) \right\} \right\rangle_{\mu_0}$$

Taking the derivative with respect to h in $h = 0$ of this equation gives us then

$$0 = \left\langle -\int_0^T dt \sum_y k(x_t, y) g(x_t, y) + \sum_{t \leq T} g(x_{t^-}, x_t) \right\rangle_{\mu_0}$$

which proves (11.3). The average of the excess entropy flux is thus

$$\langle S_{ex}(\omega) \rangle_{\mu_0} = \int_0^T dt \sum_{x,y} \mu_t(x) k(x, y) W(x, y) = \frac{1}{2} \int_0^T dt \sum_{x,y} j_{\mu_t}(x, y) W(x, y)$$

with $j_\mu(x, y) = \mu(x)k(x, y) - \mu_t(y)k(y, x)$ the probability current. This entropy flux is an excess with respect to the reference equilibrium process. For the equilibrium process we have, because of detailed balance:

$$S_0(\omega) = \log \frac{\rho(x_T)}{\rho(x_0)} = -\beta[U(x_T) - U(x_0)]$$

The average of this can be written as

$$\langle S_0(\omega) \rangle_{\mu_0} = -\beta \int_0^T \sum_x \frac{d}{dt} \mu_t(x) U(x) = \frac{\beta}{2} \int_0^T \sum_{x,y} j_{\mu_t}(x, y)[U(x) - U(y)]$$

In Section 2.8 we defined the entropy of the system through the Shannon (or Gibbs) entropy:

$$s(\mu) = -\int dx \mu(x) \log \mu(x)$$

The total average change of entropy (entropy production) in the world, given the initial distribution μ_0 is then the sum of the entropy flux of the reference process plus the excess entropy flux plus the entropy change of the system, see also (2.18):

$$\langle S_{\mu_0}(\omega) \rangle_{\mu_0} = s(\mu_T) - s(\mu_0) + \langle [S_0(\omega) + S_{ex}(\omega)] \rangle_{\mu_0} = \int_0^T dt \sigma(\mu_t) \qquad (11.4)$$

where the instantaneous entropy production rate $\sigma(\mu)$ is straightforwardly computed to be

$$\begin{aligned} \sigma(\mu) &= \frac{1}{2} \sum_{x,y} j_\mu(x, y)[\log \mu(x) - \log \mu(y) + \beta U(x) - \beta U(y) + W(x, y)] \\ &= \sum_{x,y} \mu(x) k(x, y) \log \frac{\mu(x) k(x, y)}{\mu(y) k(y, x)} \end{aligned}$$

11.1.2 Traffic

As for the entropy flux, we get the excess traffic from the action:

$$\mathcal{T}_{ex}(\omega) = A(\theta\omega) + A(\omega) = 2\int_0^T dt[\lambda(x_t) - \lambda_0(x_t)] \tag{11.5}$$

The average of this, starting from an initial distribution μ_0, is

$$\langle \mathcal{T}_{ex} \rangle_{\mu_0} = 2\int_0^T dt[\tau(\mu_t) - \tau_0(\mu_t)]$$

with

$$\tau(\mu) = \sum_x \mu(x)\lambda(x) = \sum_{x,y} \mu(x)k(x,y) \tag{11.6}$$

which is equal to the average number of jumps per unit of time for a system in a distribution μ. We call it the activity of the system. Note that this activity can be seen as the instantaneous traffic rate (only with a different reference than in (11.5)).

11.2 Observables

Like in the last chapter our first observable is the empirical time-integrated occupation density:

$$p_\omega(x) = \frac{1}{T}\int_0^T dt\delta_{x_t,x}$$

where $\delta_{x_t,x}$ is the Kronecker delta, giving one when $x_t = x$ and zero otherwise. Furthermore we define the empirical distribution that counts jumps:

$$Q_\omega(x,y) = \frac{1}{T}\sum_{t\leq T} \delta_{x_{t^-},x}\delta_{x_t,y}$$

Again, as for the case of diffusions, ergodicity leads us to the conclusion that in the long time limit $p_\omega \to \rho$ and $Q_\omega(x,y) \to \rho(x)k(x,y)$ almost surely. From Q_ω we can define the empirical current:

$$J_\omega(x,y) = Q_\omega(x,y) - Q_\omega(y,x)$$

which converges in the long time limit to j_ρ almost surely.

11.3 Joint fluctuations of occupations and currents

The easiest starting point for deriving fluctuation functionals is the joint fluctuations of the occupations and jumps:

$$P(p_\omega = \mu, Q_\omega = q) \asymp e^{-TI(\mu,q)}$$

Again, one has to take care in correctly defining such a large deviation principle, and prove when it is valid. We will not do this here. Instead we concern ourselves with computing the rate functions. We refer to [24] for mathematical details.

The heuristic argument to compute the rate function $I(\mu, q)$ is essentially the same as the one in the last chapter. We first define a new dynamics in which q and μ are typical, i.e. we take a process with transition rates $q(x, y)/\mu(x)$. Then, we compute the action for a path ω for the original process with respect to the new one:

$$A(\omega) = \int_0^T dt \sum_y [k(x_t, y) - \frac{q(x_t, y)}{\mu(x_t)}] + \sum_{t \leq T} \log \frac{q(x_{t^-}, x_t)}{\mu(x_{t^-})k(x_{t^-}, x_t)}$$

Using the definitions of our empirical observables, we rewrite this to

$$A(\omega) = T \sum_{x,y} \left\{ p_\omega(x)[k(x, y) - \frac{q(x, y)}{\mu(x)}] + Q_\omega(x, y) \log \frac{q(x, y)}{\mu(x)k(x, y)} \right\}$$

We then see that the probability of the fluctuation becomes

$$P(p_\omega = \mu, Q_\omega = q) = \int d\mathcal{P}^q(\omega) e^{-A(\omega)} \chi[p_\omega = \mu, Q_\omega = q]$$

And, similarly as for diffusions we can replace the empirical observables in the action by μ and q, so that in the end the rate function becomes

$$I(\mu, q) = \sum_{x,y} \left\{ \mu(x)k(x, y) - q(x, y) + q(x, y) \log \frac{q(x, y)}{\mu(x)k(x, y)} \right\} \tag{11.7}$$

However, we would like to find the rate function for the joint fluctuations of occupations μ and currents j. By the contraction principle, we see that this rate function is the infimum of $I(\mu, q)$ over all q for which $q(x, y) - q(y, x) = j(x, y)$. Note that for any path ω and any (bounded) state function Y, we have that

$$\sum_x Y(x) \left[\sum_y J_\omega(x, y) \right] = \frac{Y(x_0) - Y(x_T)}{T}$$

which goes to zero for the long time limit. Therefore we only consider fluctuations for which $\sum_y j(x,y) = 0$. The q^* that makes the infimum can be explicitly computed. For that we compute the Euler-Lagrange equations:

$$q(x,y) = \mu(x)k(x,y)e^{\frac{\beta F(x,y)}{2}} \tag{11.8}$$

where $F(x,y) = -F(y,x)$ is a Lagrange multiplier, which makes sure that

$$j(x,y) = q(x,y) - q(y,x) = \mu(x)k(x,y)e^{\frac{\beta F(x,y)}{2}} - \mu(y)k(y,x)e^{-\frac{\beta F(x,y)}{2}} \tag{11.9}$$

This equation can be explicitly solved, giving the minimizer q^*:

$$q^*(x,y) = \frac{j(x,y)}{2} + \frac{1}{2}\sqrt{[j(x,y)]^2 + 4\mu(x)\mu(y)k(x,y)k(y,x)}$$

This can then be substituted into $I(\mu, q = q^*)$ to get $I(\mu, j)$. Going beyond the mathematics, we compare (11.8) to the local detailed balance assumption which says that

$$\frac{q(x,y)}{q(y,x)} = \frac{\mu(x)}{\mu(y)}e^{-\beta[U(y)-U(x)-W(x,y)-F(x,y)]}$$

This gives an interpretation to the Lagrange multiplier F: it is a modification of the work $W(x,y)$ originating from a nonconservative forcing. We therefore call F a force. Moreover, the equation (11.9) tells us that this force makes the fluctuation μ, j typical. This is the same kind of modification as we had in the last chapter. There, to find the joint rate function of occupations and currents we also had to introduce a new force which made the fluctuation typical. It is illustrative to write $I(\mu, j)$ in terms of this F by substituting (11.8) into $I(\mu, q)$:

$$\begin{aligned} I(\mu,j) &= \sum_{x,y}\mu(x)k(x,y)[1 - e^{\frac{\beta}{2}F(x,y)}] + \frac{\beta}{4}\sum_{x,y} j(x,y)F(x,y) \\ &= \tau(\mu) - \tau_F(\mu) + \frac{\beta}{4}\sum_{x,y} j(x,y)F(x,y) \end{aligned} \tag{11.10}$$

In the last line we have written the excess activity of the system, excess meaning the modified dynamics with respect to the original one. Similarly, the last term in the rate function is an excess entropy flux, given a current j. This last expression coincides physically with (10.25), and is as such the most general and physical form of writing the rate function.

We can also use the reference equilibrium process to rewrite the fluctuation functional into a form even more clearly similar to (10.25):

$$\begin{aligned} I(\mu, j) &= \sum_{x,y} \mu(x) k_0(x,y) [e^{\frac{\beta}{2} W(x,y)} - e^{\frac{\beta}{2} G(x,y)}] \\ &\quad + \frac{\beta}{4} \sum_{x,y} j(x,y)[G(x,y) - W(x,y)] \\ &= \tau_W(\mu) - \tau_G(\mu) + \frac{\beta}{4} \sum_{x,y} j(x,y)[G(x,y) - W(x,y)] \end{aligned}$$

where $G = W + F$ is the nonconservative forcing in the modified dynamics, and we have denote subscripts W and G to denote the traffic in the original and modified dynamics.

11.4 A notion of thermodynamic potentials

Like in Section 10.5, we can use the traffic of a process to construct a potential. The activity, as defined in (11.6), has the following property: define a dynamics with rates $k_h(x,y) = k_0(x,y) \exp\{\frac{\beta}{2} h(x,y)\}$, where $h(x,y) = -h(y,x)$. then the function $\mathcal{H}(\mu, h) = \tau_h(\mu) - \tau_0(\mu)$ satisfies

$$\frac{\delta \mathcal{H}(\mu, h)}{\delta h(x,y)} = \frac{\beta}{2} [\mu(x) k_h(x,y) - \mu(y) k_h(y,x)] = \frac{\beta}{2} j_\mu^h(x,y)$$

meaning that the activity (traffic) is a potential for the currents, like the traffic was in the overdamped diffusion case. Analogous to that case we can define a Legendre transform:

$$G(\mu, j) = \sup_h [\frac{\beta}{4} \sum_{x,y} j(x,y) h(x,y) - \mathcal{H}(\mu, h)]$$

This is a potential for forces: the minimizer h is found to be the force $G = W + F$ that makes $j = j_\mu^G$ typical, and

$$\frac{\delta G}{\delta j(x,y)} = \frac{\beta}{2} G(x,y)$$

The fluctuation functional $I(\mu, j)$ can then be written as

$$I(\mu, j) = \mathcal{H}(\mu, W) + G(\mu, j) - \frac{\beta}{4} \sum_{x,y} j(x, y) W(x, y)$$

Exactly like for the diffusion case (10.30). Again, $G(\mu, j)$ is equal to $I_{W=0}(\mu, j)$, i.e. the equilibrium fluctuation functional.

11.5 Contraction to occupations

To find the fluctuation functional for occupations alone, we make the contraction

$$I(\mu) = \inf_j I(\mu, j) \quad \text{or} \quad I(\mu) = \inf_q I(\mu, q)$$

The infimum has to be taken with the restriction that $\sum_y [q(x, y) - q(y, x)] = \sum_y j(x, y) = 0$. If we do the contraction using $I(\mu, q)$, we get Euler-Lagrange equations of the following form:

$$\begin{aligned} q(x, y) &= \mu(x) k(x, y) e^{\frac{\beta}{2}[V(y) - V(x)]} \\ 0 &= \sum_y \left[\mu(x) k(x, y) e^{\frac{\beta}{2}[V(y) - V(x)]} - \mu(y) k(y, x) e^{\frac{-\beta}{2}[V(y) - V(x)]} \right] \end{aligned}$$

where V is a Lagrange multiplier making sure that $\sum_y [q(x, y) - q(y, x)] = \sum_y j(x, y) = 0$. Although we can not solve this in general, we see that V is a potential to be added to the dynamics that makes μ typical. This is analogous to the overdamped diffusion case. Substituting this form of q in $I(\mu, q)$ we can write $I(\mu)$ in terms of this potential:

$$I(\mu) = \sum_{x,y} \left\{ \mu(x) k(x, y) - \mu(x) k(x, y) e^{\frac{\beta}{2}[V(y) - V(x)]} \right\} = \tau(\mu) - \tau_V(\mu) \quad (11.11)$$

Indeed, the fluctuation functional for occupations alone is defined in terms of an excess traffic, just as for the case of overdamped diffusions. In contrast to that case, however, this excess is not equal to an excess of entropy production. It seems therefore, that the excess traffic is a more general quantity than entropy to characterize fluctuation functionals.

11.6 Small fluctuations

Consider the regime of small fluctuations, i.e. where the quadratic approximation of the fluctuation functionals is valid. We write $\mu(x) = \rho(x)[1 + \epsilon\mu_1(x)]$ and $j(x,y) = j_\rho(x,y) + \epsilon j_1(x,y)$ with $\epsilon \in \mathbb{R}$ a small number. Note that j_1 has to satisfy the constraint $\sum_y j_1(x,y) = 0$. To write down $I(\mu, j)$ up to second order in ϵ, we expand first (11.9), giving $F(x,y) = \epsilon F_1(x,y) + o(\epsilon)$, with

$$F_1(x,y) = \frac{1}{\tau_\rho(x,y)}\Big[j_1(x,y) - \frac{1}{2}\tau_\rho(x,y)[\mu_1(x) - \mu_1(y)] - \frac{1}{2}j_\rho(x,y)[\mu_1(x) + \mu_1(y)]\Big]$$

where we have defined the symmetric counterpart of the current:

$$\tau_\mu(x,y) = \mu(x)k(x,y) + \mu(y)k(y,x)$$

which is the expected number of jumps between states x and y per unit of time, when the system is in a distribution μ. This quantity is closely related to the activity defined in (11.6), because

$$\sum_{x,y} \tau_\mu(x,y) = 2\tau(\mu)$$

The factor of two in this equality stems from the fact that every bond x, y is counted twice in the sum. We therefore also call $\tau_\mu(x,y)$ traffic (or activity). The fluctuation functional $I(\mu, j)$ in terms of the added force F becomes

$$I(\mu, j) = \frac{\beta^2\epsilon^2}{8}\sum_{x,y} \tau_\rho(x,y)[F_1(x,y)]^2 + o(\epsilon^2)$$

and substituting the explicit form of F, we get schematically:

$$I(\mu, j) = \frac{\beta^2\epsilon^2}{8}\sum_{x,y} \frac{1}{\tau_\rho}[j_1 - \tau_\rho\nabla^-\mu_1 - j_\rho\nabla^+\mu_1]^2(x,y) + o(\epsilon^2) \tag{11.12}$$

with the notation $\nabla^\pm\mu_1(x,y) = \frac{1}{2}[\mu_1(x) \pm \mu_1(y)]$. In this quadratic approximation, we see that the stationary traffic τ_ρ plays the role of a variance.

We can write (11.12) in the form

$$I(\mu, j) = \frac{\beta^2\epsilon^2}{4}\sum_{x,y}\Big[\frac{j_1^2}{2\tau_\rho} + \frac{\tau_\rho}{2}(\nabla^-\mu_1)^2 - \frac{j_\rho j_1}{\tau_\rho}\nabla^+\mu_1 + \frac{j_\rho^2}{2\tau_\rho}(\nabla^+\mu_1)^2\Big](x,y)$$
$$+o(\epsilon^2) \tag{11.13}$$

which demonstrates that there is an occupation-current coupling, which is proportional to the stationary current. It is this coupling that makes contractions to occupation densities or currents alone difficult.

The close-to-equilibrium regime Let us assume that the dynamics of the process is close to equilibrium, meaning that the transition rates are of the form (11.1):

$$k(x,y) = k_0(x,y)e^{\frac{\beta\epsilon}{2}W(x,y)}$$

with $k_0(x,y)$ the equilibrium reference process. In such a dynamics, the stationary current $j_\rho = O(\epsilon)$. In (11.13) we can therefore drop the terms containing the stationary current to obtain

$$I(\mu, j) = \frac{\beta^2\epsilon^2}{4}\sum_{x,y}\Big[\frac{j_1^2}{2\tau_\rho} + \frac{\tau_\rho}{2}(\nabla^-\mu_1)^2\Big](x,y) + o(\epsilon^2) \tag{11.14}$$

which gives us the same conclusions as for the case of overdamped diffusions: the current and occupation fluctuations decouple in this close-to-equilibrium regime. As a consequence a minimum entropy production principle can be derived for the occupation fluctuations and a maximum entropy production principle for currents.

A question arises here especially for the occupation fluctuations, which was not visible for overdamped diffusions: why does entropy production govern the close-to-equilibrium regime? The occupation fluctuation functional is expressed as an excess traffic. For overdamped diffusions this was equal to an excess entropy production, but here it is not.

However, in the close-to-equilibrium regime, the excess traffic and excess entropy production merge. We see this by inserting $\mu = \rho_0[1 + \epsilon\mu_1]$ and $k(x,y) = k_0(x,y)e^{\frac{\beta\epsilon}{2}W(x,y)}$ into the definitions of traffic and entropy production. Note that ρ_0 is the equilibrium distribution for the rates k_0. Using subscripts W to denote that we work in the dynamics with W we get

$$\sigma_W(\mu) = \frac{\epsilon^2}{2}\sum_{x,y}\rho_0(x)k_0(x,y)\Big[\mu_1(x) - \mu_1(y) + W(x,y)\Big]^2 + o(\epsilon^2)$$

$$\tau_W(\mu) = \sum_{x,y} \rho_0(x)k_0(x,y)\Big[1+\epsilon\mu_1(x)+\frac{\epsilon}{2}W(x,y) + \frac{\epsilon^2}{4}[\mu_1(x)-\mu_1(y)]W(x,y)+\frac{\epsilon^2}{8}W(x,y)^2\Big] + o(\epsilon^2)$$

Up to second order in ϵ, the excess entropy production and traffic with respect to the equilibrium dynamics thus give the same:

$$\sigma_W(\mu) - \sigma_0(\mu) = \tau_W(\mu) - \tau_0(\mu) + o(\epsilon^2) \tag{11.15}$$

The reason that entropy production principles appear close to equilibrium is thus a result of two facts: The first is that the occupation and current fluctuations decouple, so that the joint fluctuation functional becomes a sum of the functionals for occupations and currents alone. Second, the traffic and entropy production become indistinguishable, so that we do not need traffic do describe these physical results.

Chapter 12

Conclusions

The questions and the methods covered in this part are not entirely original. They have appeared in the mathematical literature in a systematic way since the theory of large deviations was introduced in the framework of Markov processes, see the references in Chapter 9 and especially in Section 9.6. The relevance to physics and to statistical mechanics in particular is obvious, but the thermodynamic interpretation of the resulting dynamical fluctuation functionals has not been systematically investigated. A first study can be found in [69]. The research discussed in this part and in [72, 96] has added to that. There have of course been many other studies of dynamical fluctuation theory in the literature.

We mention in particular the works of Derrida and Bodineau, see [9, 25] and of Bertini *et al*, e.g. in [5, 6]. With respect to diffusion processes, our approach is especially similar to what one is doing for a macrostatistical theory where a hydrodynamic limit is taken by rescaling of certain Markov jump processes. The hydrodynamic fluctuations can then be viewed as a solution of some infinite-dimensional diffusion process. Possible differences with the existing work are first of all that the problems related to the diffusion-approximation or to a hydrodynamic rescaling do not enter in our work. We just start from a finite-dimensional diffusion process or a Markov jump process as such, and without extra rescaling. In other words, we prefer to split the problem of hydrodynamical scaling with possible diffusion approximation from the problem of studying the dynamical fluctuations.

The emphasis of our work is on the distinction between the time-symmetric and time-antisymmetric sector of the fluctuations. On the one hand the action of relative path-probability densities is split in entropy flux and traffic. On the other hand the observables are split into occupation density (time-symmetric) and current (time-antisymmetric). It then comes as no surprise that the joint

fluctuation functional $I(\mu, j)$ can be written as an excess traffic, and an excess entropy flux, which constitutes a first main conclusion, expressed in (10.25) and (11.10). The excess is an excess of the original process with respect to a modified process (by adding a force to the dynamics) in which the fluctuations are typical. Contractions to occupations or currents alone is hard to do.

Our second main result covers occupation fluctuations, see (10.34) and (11.11). There one can see that the fluctuation functional can be expressed in terms of only an excess traffic, and the modified dynamics is made by adding a potential. Furthermore, we like to emphasize a structure in the joint fluctuations of density and current which reminds of thermodynamic potentials, which can be found in Sections 10.5 and 11.4.

Finally, for small fluctuations one can clearly see that the coupling between occupation and current fluctuations only vanishes when close to equilibrium. In that regime minimum and maximum entropy production principles emerge.

In [72] we have investigated dynamical fluctuations for semi-Markov jump processes. These are an extension of Markov jump processes where the waiting times are not exponentially distributed. In [96] we extended the analysis of overdamped diffusion processes to time-dependent but periodic dynamics. In those cases the general conclusions we have outlined do remain true.

However, it is clear that the story as presented in this part is far from finished. First of all, we have not provided immediately applicable results. We have rather begun to examine the general physical structure of dynamical fluctuations. To go on, other models should be studied to see which structures remain in more general settings. For example, one could see what happens for underdamped diffusions, as we did for fluctuation-dissipation relations. Large deviations in this setting have already been discussed in [10], where a quantity similar to our traffic plays a major role.

Part IV

Conclusions and appendices

"While working on my note, I realized that I had met these guys (=traffic) before, when I had been trying to prove some of my conjectures on the driven lattice gas. Those conjectures came mainly from my 'intuition' that was unfortunately based on the equilibrium physics, and turned out to be wrong. Always it looked like the 'proof' went OK, but at some stage I saw something went wrong because of unwanted factors coming from waiting times. That was traffic, and (in those days) I only regarded it as annoying noise. But your message probably is to make friends with them. I will try."

Hal Tasaki, Department of Physics, Gakushuin University, 171-8588, JAPAN, in a recent e-mail conversation (September 2009)

Chapter 13

Overall conclusions

This chapter is devoted to some general conclusions of this text. For conclusions about fluctuation-dissipation relations and dynamical fluctuations separately, we refer to Chapters 8 and 12. Basically we have used the same method throughout this text: on the level of stochastic trajectories we distinguish entropy and traffic, that completely specify the probabilities of those trajectories. Quantities and physical relations that are computed through averages over possible trajectories are thus governed by both entropy and traffic. Entropy is a known quantity: it has an operational definition. In this chapter we shortly discuss the status of the less-known traffic and its possible interpretations.

13.1 Traffic for Markov jump processes and overdamped diffusions

We have investigated both fluctuation dissipation-relations and dynamical fluctuations for Markov jump processes and overdamped diffusions. This warrants a comparison. Generally we have used the same strategy in both parts, namely using a description in terms of stochastic trajectories and more importantly the splitting of the action in a time-symmetric and time-antisymmetric part.

The time-antisymmetric part of the action is, under the local detailed balance assumption, equal to the entropy flux into the environment. The restrictions under which this works is that the parts of the environment connected to the system, each are and remain in equilibrium. This entropy flux can therefore be expressed in terms of the heat dissipation into that environment.

The time-symmetric part of the action is called traffic. It is much more difficult to give this quantity a general physical meaning than for entropy flux. One reason for that, is that the concept only makes sense as an excess. This excess is a difference of the same quantity in two dynamics. Our results for Markov jump processes and overdamped diffusions do give us a tool to give more interpretation to traffic.

13.1.1 Traffic as an escape rate

We have seen that the rate function for the large deviations of the empirical occupation density $I(\mu)$ can be expressed as an excess in traffic (a traffic rate actually), see (10.34) and (11.11).

How can we interpret this rate function? For this, first consider the set of distributions μ that vanish on a given subset A of the state space Ω. The contraction principle then tells us that the infimum of $I(\mu)$ over this set gives the probability of the particle not visiting this A. This infimum can thus be considered an escape rate from the subset A. The rate function $I(\mu)$, and thus the excess traffic can then be called a generalized escape rate from the distribution μ.

Another way of seeing this is by reversing our viewpoint. The original viewpoint is that, given a distribution μ we look for a potential V that makes it typical. Then we compute the excess traffic of the original with respect to this new dynamics with the added potential. The reverse viewpoint is that we start from some dynamics with a stationary density μ. We then add to the dynamics a potential $-V$. The change in the instantaneous traffic rate then gives the rate function of μ in that new dynamics, in which it is no longer stationary. The bigger the rate function (excess traffic) the faster μ relaxes to the new stationary distribution, thus 'escaping from μ.' An interesting observation follows from this reverse viewpoint: when a potential is added to a stationary system, the positivity of the rate function tells us that the traffic should increase. In what follows, we interpret traffic as activity, meaning that adding a potential increases the activity of the system.

Also on the level of trajectories we can see this. For Markov jump processes we wrote (11.5), which is a time-integral of the escape rates of the states the system has visited. Moreover, overdamped diffusions can be recovered from Markov jump processes when space and time are rescaled in a specific way (see Appendix C). Therefore we still interpret the traffic in terms of 'excess escape rates' even on the level of paths for overdamped diffusions.

13.1.2 Traffic as activity

The excess traffic appearing in the rate function for the large deviations of the empirical occupation density $I(\mu)$ is, for Markov jump processes, equal to

$$I(\mu) = \tau(\mu) - \tau_V(\mu) = \sum_{x,y} \mu(x)k(x,y) - \sum_{x,y} \mu(x)k(x,y)e^{\frac{\beta}{2}[V(y)-V(x)]}$$

This is an excess in the expected number of jumps per unit of time, when in the distribution μ. In this sense, the traffic can be called an activity. Again, because overdamped diffusions are recovered as a scaling limit of Markov jump processes, we can still see the excess traffic as an excess activity for overdamped diffusion processes.

When the fluctuation is small, i.e. μ is close to the stationary distribution, the potential V that makes μ typical, is also small. if we write $\mu = \rho[1 + \epsilon\mu_1]$ and $V \to \epsilon V$, we get:

$$I(\mu) = \tau(\mu) - \tau_V(\mu) = -\frac{\beta\epsilon^2}{2} \sum_{x,y} \rho(x)\mu_1(x)k(x,y)[V(y) - V(x)] \quad (13.1)$$

$$= -\frac{\beta\epsilon^2}{2} \sum_{x} \rho(x)\mu_1(x)LV(x) \quad (13.2)$$

and similarly for overdamped diffusions. Indeed, we recognize here the version of traffic $\tau(x) = \beta LV(x)$ that appeared in the fluctuation-dissipation relations (6.6).

13.1.3 Unpopular traffic

When comparing the fluctuation-dissipation relation out of equilibrium, but for one heat bath (5.14) with the equilibrium one (5.2), we see that the equilibrium fluctuation-dissipation theorem can be rewritten:

$$R_{QV}(t,s) = \beta \frac{\partial}{\partial s} \langle V(x_s)Q(x_t)\rangle^0_{\mu_0} = -\langle \tau(\omega,s)Q(x_t)\rangle^0_{\mu_0}$$

The first equality is the usual one. In our framework we see that it is the correlation between the observable and the excess entropy flux of the perturbed with respect to the unperturbed process. The last equality states that we can also consider the correlation with the excess traffic. In equilibrium these are the same!

Similarly, for dynamical fluctuations we have seen that close to equilibrium excess entropy production and excess traffic are the same, see (11.15). This may

explain why traffic has not appeared much earlier in statistical physics. Around equilibrium one does not see it as an independent new quantity, and all attention goes to its more popular brother entropy. This explains results as minimum and maximum entropy production principles, (see Sections 10.7 and 11.6) and the McLennan formula (Section 2.10).

Celebrated recent results concerning a far-from-equilibrium regime, like the fluctuation-theorem and work relations (see Section 2.10), only make statements about the time-antisymmetric part of the dynamics. Because of that entropy is sufficient to describe the results, avoiding traffic completely.

It seems however, that for a full description of nonequilibrium systems, the time-symmetric sector of the process described by the traffic is indispensable. This constitutes a major message of this text: traffic should not be avoided, it should be investigated.

13.2 Outlook

Fluctuation-dissipation relations In part II of this text we have outlined the general strategy to derive fluctuation-dissipation relations out of equilibrium, as well as given several explicit examples, ready to be tested experimentally. But of course the study is not finished. Apart from applying the strategy to other models, several related problems seem very interesting and can be investigated in the same framework. We suggest some examples:

- The response of a system with respect to a perturbation other than a potential. This could be for example an added nonconservative force, or a change in the chemical potential of one of the particle reservoirs. This should be perfectly doable within the framework outline in this text.
- Changes in temperature of the heat baths as perturbations. This is a very interesting topic for applications in meteorology but also in nonequilibrium calorimetry. However, such a perturbation can be tricky in our framework: for example for diffusion processes: if the diffusion coefficient (which depends usually on the temperature) is perturbed, then the path-probability densities of the perturbed and unperturbed process are no longer absolutely continuous with respect to each other.
- A step towards thermodynamics through quasistatic processes: suppose that amplitude h_t of the perturbation potential is not small, but changes very slowly. So slowly that the system is always very close to the stationary distribution of the dynamics at that moment. In such cases the response relations we have derived can be used to write expressions for the excess

heat, work and change of energy. The second law of thermodynamics for example can be investigated in such processes.

- The use of our results in models of spin-glasses, where the notion of effective temperature often is useful. As our approach gives the response explicitly in terms of correlation functions for Markov jump processes, which often model spin-glasses, it could be used to develop the knowledge of effective temperatures.

Dynamical fluctuations In part III we have discussed dynamical fluctuations of occupations and currents in overdamped diffusions and Markov jump processes. Although some quite general conclusions were made in terms of entropy and traffic, and this theory adds physical interpretation to traffic as escape rate and as a 'thermodynamic' potential, the story is far from finished. Some suggestions for future work:

- An investigation of the connection between dynamical and static fluctuations, how one theory can be derived from the other, can give a more systematic understanding of large deviations in nonequilibrium statistical mechanics.
- Dynamical fluctuations in underdamped diffusion processes can give more insight into the concept of traffic in these models. As is clear from Markov jump processes and overdamped diffusions, dynamical fluctuation theory is of great help here.
- It would be good to have some explicit results, i.e. some explicit rate functions, which moreover relate to experimental accessible predictions. Simulations could also help in visualizing the rate functions, at least for simple systems.

Appendix A

The generator for Markov processes

A.1 Different generators

Backward generator When the system evolves according to a Markov dynamics, there exists an operator L acting on state functions f, such that

$$\frac{d}{dt}\left\langle f(x_t)\right\rangle_{\mu_0} = \left\langle Lf(x_t)\right\rangle_{\mu_0} \qquad \forall f \tag{A.1}$$

This is a differential equation with the following solution

$$\left\langle f(x_t)\right\rangle_{\mu_0} = \int dx \mu_0(x)(e^{tL}f)(x)$$

where the integral should be replaced by a sum when the configuration space is discrete. In words, e^{tL} "pulls a function f back to the time of the initial measure." L is therefore often called the backward generator.

Forward generator There also exists a forward generator $L^\dagger$, defined in the following way: for any pair of state functions f, g

$$\int dx g(x) Lf(x) = \int dx f(x) L^\dagger g(x) \tag{A.2}$$

If we now consider the definition of the time-evolved distribution μ_t:

$$\int dx \mu_0(x)(e^{tL}f)(x) = \langle f(x_t)\rangle_{\mu_0} = \int dx \mu_t(x) f(x)$$

we see that $\mu_t = e^{tL^\dagger}\mu_0$. So $L^\dagger$ pushes the distribution μ_0 forward in time, and is therefore called the forward generator. The forward generator can be found from the Master equation for Markov jump processes or the Fokker-Planck equation for diffusions, because

$$\frac{\partial \mu_t}{\partial t} = L^\dagger \mu_t \tag{A.3}$$

Adjoint generator For time-independent dynamics, the adjoint generator L^* is defined with the help of the stationary distribution ρ: for any pair of state functions f, g

$$\int dx \rho(x) g(x) L^* f(x) = \int dx \rho(x) f(x) L g(x) \tag{A.4}$$

It is very important to realize that this adjoint generator depends on the stationary distribution, and is therefore often not explicitly known in nonequilibrium systems. This adjoint generator is related to the time-reversed dynamics.

To apply time-reversal, we must remember that there are possibly variables that change sign under time-reversal, i.e. velocities. To simplify the analysis however, we restrict ourselves here to the case that we have only time-symmetric variables (Markov jump processes and overdamped diffusions).

The stationary distribution for the time-reversed dynamics (with path-probability density $\mathcal{P}(\theta\omega)$) is also ρ, and for any f, g:

$$\begin{aligned}
\langle g(x_0) f(x_t)\rangle_\rho &= \int d\mathcal{P}_\rho(\omega) g(x_0) f(x_t) \\
&= \int d\mathcal{P}_\rho(\theta\omega) g(x_t) f(x_0) \\
&= \langle g(x_t) f(x_0)\rangle_\rho^\theta
\end{aligned}$$

where the θ denotes that the average is taken in the time-reversed dynamics. Writing the generator of the time-reversed dynamics as L^θ, we get

$$\int dx\rho(x)g(x)e^{tL}f(x) = \int dx\rho(x)f(x)e^{tL^\theta}g(x)$$

Because of the definition (A.4), we see that $L^\theta = L^*$. More generally, one can derive that $L^\theta = \pi L^*\pi$ in the case that the configuration of the system contains velocities. In other words, up to some reversal of signs of velocities, the adjoint generator is the generator of the time-reversed dynamics. This also means that an equilibrium dynamics must satisfy $L^* = \pi L\pi$.

Here we also see that the equilibrium distribution can be defined as the distribution that satisfies

$$\int dx\rho(x)f(x)Lg(x) = \int dx\rho(x)g(x)\pi L\pi f(x) \qquad \forall f, g \tag{A.5}$$

A.2 Correlation functions

With the definitions of generators, one can write correlation functions in several different ways as follows, for $s \leq t$

$$\begin{aligned}
\langle f(x_t)g(x_s)\rangle_{\mu_0} &= \int dx\mu_0(x)e^{sL}(ge^{(t-s)L}f)(x)dx \\
&= \int dxg(x)e^{(t-s)L}f(x)e^{sL^\dagger}\mu_0(x)dx \\
&= \int dx\mu_s(x)g(x)e^{(t-s)L}f(x)dx \\
&= \langle f(x_{t-s})g(x_0)\rangle_{\mu_s}
\end{aligned}$$

One can see from this, that the time derivative of such a correlation function can be quite complicated. For example, for $s < t$:

$$\begin{aligned}
\frac{d}{ds}\langle f(x_t)g(x_s)\rangle_{\mu_0} &= -\langle Lf(x_t)g(x_s)\rangle_{\mu_0} + \left\langle f(x_t)\frac{g(x_s)}{\mu_s(x_s)}\frac{\partial\mu_s}{\partial s}(x_s)\right\rangle_{\mu_0} \\
&= -\left\langle f(x_t)\frac{L^\dagger(\mu_s g)}{\mu_s}(x_s)\right\rangle_{\mu_0} + \left\langle f(x_t)\frac{g(x_s)}{\mu_s(x_s)}L^\dagger\mu_s(x_s)\right\rangle_{\mu_0}
\end{aligned}$$

$$\frac{d}{dt}\left\langle f(x_t)g(x_s)\right\rangle_{\mu_0} = \left\langle Lf(x_t)g(x_s)\right\rangle_{\mu_0} \tag{A.6}$$

However, we see that in stationary averages, this simplifies to

$$\frac{d}{ds}\left\langle f(x_t)g(x_s)\right\rangle_{\rho} = -\left\langle Lf(x_t)g(x_s)\right\rangle_{\rho} = -\frac{d}{dt}\left\langle f(x_t)g(x_s)\right\rangle_{\rho}$$

When the generator occurs in a stationary correlation function, one can replace it by the adjoint generator in the following case, for $s \leq t$:

$$\begin{aligned}
\left\langle Lf(x_t)g(x_s)\right\rangle_{\rho} &= \int dx\rho(x)g(x)e^{(t-s)L}Lf(x)dx \\
&= \int dx\rho(x)g(x)Le^{(t-s)L}f(x)dx \\
&= \int dx\rho(x)L^*g(x)e^{(t-s)L}f(x)dx \\
&= \left\langle f(x_t)L^*g(x_s)\right\rangle_{\rho}
\end{aligned}$$

where we used in the second equality that L commutes with $e^{(t-s)L}$. However, we can not apply a similar procedure for $\left\langle f(x_t)Lg(x_s)\right\rangle_{\rho}$, because L and L^* do not commute in general.

A.3 Generators for jump processes and diffusions

As additional information, we give here the generators for the main models used in this thesis. The procedure to find them is straightforward. We start from the evolution equation for the distribution μ_t, which is the Master equation for Markov jump processes and the Fokker-Planck equation for diffusions, to find the forward generator $L^\dagger$:

$$\frac{\partial \mu_t}{\partial t} = L^\dagger \mu_t$$

After that we use (A.2) to find the backward generator, and (A.4) to find the adjoint generator.

A.3.1 Markov jump processes

The generators for a general Markov jump process with transition rates $k(x, y)$ are:

$$L^\dagger \mu(x) = \sum_y [\mu(y)k(y,x) - \mu(x)k(x,y)] \tag{A.7}$$

$$Lf(x) = \sum_y k(x,y)[f(y) - f(x)] \tag{A.8}$$

$$L^* f(x) = \sum_y \frac{\rho(y)k(y,x)}{\rho(x)}[f(y) - f(x)] \tag{A.9}$$

A.3.2 Overdamped diffusions in more dimensions

The generators for a general overdamped Langevin dynamics (4.13) are:

$$L^\dagger \mu = -\nabla \cdot [\chi F \mu - D\nabla\mu] \tag{A.10}$$

$$Lf = F\chi\nabla f + \nabla \cdot (D\nabla f) \tag{A.11}$$

$$L^* f = -F\chi\nabla f - \nabla \cdot (D\nabla f) + 2\frac{\nabla\rho}{\rho} D\nabla f \tag{A.12}$$

A.3.3 Underdamped diffusions

The generators for underdamped diffusions of the form (7.1) are:

$$L^\dagger \mu = \sum_i \left[-v_i \frac{\partial \mu}{\partial x_i} - \frac{\partial}{\partial v_i}[(f_i - \frac{\partial U}{\partial x_i} - m_i\gamma_i v_i)\mu] + D_i \frac{\partial^2 \mu}{\partial v_i^2} \right] \tag{A.13}$$

$$Lg = \sum_i \left[v_i \frac{\partial g}{\partial x_i} + (f_i - \frac{\partial U}{\partial x_i} - m_i\gamma_i v_i)\frac{\partial g}{\partial v_i} + D_i \frac{\partial^2 g}{\partial v_i^2} \right] \tag{A.14}$$

$$L^* g = \sum_i \left[-v_i \frac{\partial g}{\partial x_i} - (f_i - \frac{\partial U}{\partial x_i} - m_i\gamma_i v_i)\frac{\partial g}{\partial v_i} + D_i \frac{\partial^2 g}{\partial v_i^2} - 2D_i \frac{1}{\rho}\frac{\partial \rho}{\partial v_i}\frac{\partial g}{\partial v_i} \right]$$

Appendix B

A derivation of the Fokker-Planck equation

Starting from a Langevin equation it is possible to derive the corresponding Fokker-Planck equation. We show it here for an overdamped Langevin equation in one dimension (4.10):

$$dx_t = \chi F(x_t)dt + \sqrt{2D}dB_t$$

We do this by deriving the generator L of this process. For this, take an arbitrary state function g, which is smooth. Take a small number $\epsilon \in \mathbb{R}$. We consider the average

$$\langle g(x_{t+\epsilon})\rangle_{\mu_0} = \int_{\mathbb{R}} dx \mu_t(x) \int_{\mathbb{R}} dy Prob(x_{t+\epsilon} = x + y | x_t = x) g(x + y)$$

where, for ϵ small enough we can write

$$Prob(x_{t+\epsilon} = x + y | x_t = x) = \frac{1}{N} \exp\left\{ -\frac{1}{4D\epsilon}[y - \chi F(x)\epsilon]^2 \right\}$$

with the normalization $N = \sqrt{4\pi D\epsilon}$. We make a Taylor-expansion of g around x:

$$\int_{\mathbb{R}} dy Prob(x_{t+\epsilon} = x + y | x_t = x) g(x+y)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n g}{\partial x^n}(x) \int_{\mathbb{R}} dy \frac{1}{N} e^{-\frac{1}{4D\epsilon}[y - \chi F(x)\epsilon]^2} y^n$$

In this way we have reduced the problem to calculating Gaussian integrals. We find

$$\begin{aligned} \int_{\mathbb{R}} dy Prob(x_{t+\epsilon} = x + y | x_t = x) g(x+y) &= g(x) + \epsilon \chi F(x) \frac{\partial g}{\partial x}(x) \\ &\quad + \epsilon D \frac{\partial^2 g}{\partial x^2}(x) + o(\epsilon) \end{aligned}$$

So that in the limit of $\epsilon \to 0$ we get

$$\begin{aligned} \langle Lg(x_t) \rangle_{\mu_0} &= \lim_{\epsilon \to 0} \frac{\langle g(x_{t+\epsilon}) \rangle_{\mu_0} - \langle g(x_t) \rangle_{\mu_0}}{\epsilon} \\ &= \int_{\mathbb{R}} dx \mu_t(x) [\chi F(x) \frac{\partial g}{\partial x}(x) + D \frac{\partial^2 g}{\partial x^2}(x)] \end{aligned}$$

This gives us the backward generator L. The forward generator $L^\dagger$ then defines the Fokker-Planck equation. It is easily found by (A.2).

The generators and Fokker-Planck equations for other and more general diffusion equations can be found by the same reasoning.

Appendix C

From Markov jump processes to overdamped diffusions

It is possible to derive overdamped diffusions as a certain limit of Markov jump processes. As an example, we consider the case of a single particle performing a Markov jump process on a ring with N sites, labelled $x = 0, \epsilon, 2\epsilon, \ldots, N\epsilon$ with $N\epsilon = 1 \equiv 0$. The particle can hop every time only one site to the left or to the right, with rates $k(x, x \pm \epsilon)$. A Markov process is fully determined by its generator. We therefore prove that the generator of this jump process converges in a certain scaling for $\epsilon \to 0$ to the generator of a diffusion process. The generator of the Markov jump process on a state function g gives:

$$\begin{aligned} L_\epsilon g(x) &= \sum_{y=x\pm\epsilon} k(x, x \pm \epsilon)[g(x \pm \epsilon) - g(x)] \\ &= \frac{1}{2}[k(x, x+\epsilon) + k(x, x-\epsilon)][g(x+\epsilon) + g(x-\epsilon) - 2g(x)] \\ &\quad + \frac{1}{2}[k(x, x+\epsilon) - k(x, x-\epsilon)][g(x+\epsilon) - g(x-\epsilon)] \end{aligned}$$

We expand the transition rates in orders of ϵ and rescale them as follows with $\epsilon \to 0$:

$$k(x, x+\epsilon) = \frac{1}{\epsilon^2}[k_0(x) + \epsilon k_1(x) + o(\epsilon)]$$

for the rates for hopping to the right. This corresponds to a scaling of space (the space between two sites is $1/\epsilon$) and of time (the escape rates of the process are

multiplied by $1/\epsilon^2$ so that the process goes faster and faster).

Furthermore, the local detailed balance condition tells us that

$$\frac{k(x,x+\epsilon)}{k(x+\epsilon,x)} = \exp\{\beta[W(x,x+\epsilon)+U(x)-U(x+\epsilon)]\}$$

where $W(x,x+\epsilon)$ is the work of the nonconservative forcing when the particle travels a distance ϵ to the right. We write it therefore as $W(x,x+\epsilon)=\epsilon f(x)+o(\epsilon)$. This allows us to express the rates for hopping to the left as:

$$\begin{aligned} k(x+\epsilon,x) &= k(x,x+\epsilon)\exp\{-\beta\epsilon f(x)+\beta\epsilon U'(x)+o(\epsilon)\} \\ &= \frac{1}{\epsilon^2}\left[k_0(x)+\epsilon[k_1(x)-\beta k_0(x)f(x)+\beta k_0(x)U'(x)]+o(\epsilon)\right] \end{aligned}$$

For our calculation, we need $k(x,x-\epsilon)$, which is thus equal to

$$\frac{1}{\epsilon^2}\left[k_0(x)+\epsilon[k_1(x)-k_0'(x)-\beta k_0(x)f(x)+\beta k_0(x)U'(x)]+o(\epsilon)\right]$$

Substituting this into the generator, we get the following limit:

$$\lim_{\epsilon\to 0} L_\epsilon g(x) = k_0(x)g''(x)+[\beta k_0(x)f(x)-\beta k_0(x)U'(x)+k_0'(x)]g'(x)$$

which exactly defines the generator for a diffusion process with the Langevin equation

$$dx_t = \chi(x)[f(x_t)-U'(x_t)]dt + D'(x_t)dt + \sqrt{2D(x_t)}dB_t$$

with $D=k_0$ and $\chi=\beta k_0$. We therefore see the forcing with D' naturally emerging. Moreover, the local detailed balance condition for Markov jump processes, has the relation $\chi=\beta D$ as a consequence after the rescaling.

Appendix D

Samenvatting

Hier volgt een samenvatting van het onderzoek beschreven in deze thesis, in het Nederlands. Eerst geven we echter een algemene inleiding, bedoeld voor een breed publiek.

D.1 Inleiding

Onze dagelijkse wereld bestaat uit een enorm aantal deeltjes (atomen en moleculen). Een gewone kop koffie bijvoorbeeld, heeft er zo'n 10^{23}. Wanneer we over één of twee deeltjes praten, dan kunnen we heel precies berekenen hoe die gaan bewegen. Dan zijn we bezig met 'mechanica,' en dat doet men in de fysica met de wetten van Newton. Maar 10^{23}, dat zijn zoveel deeltjes, dat niemand kan voorspellen hoe die allemaal exact gaan bewegen. Dat is ook niet erg, we hoeven dat ook niet te weten. We zijn eerder geïnteresseerd in een aantal 'bruikbare' eigenschappen, zoals temperatuur, volume en energie. Zo kan het handig zijn om te weten hoe snel een kop koffie afkoelt of hoeveel brandstof je nodig hebt om van Leuven naar Brussel te rijden. Meer algemeen zijn we in dit soort situaties geïnteresseerd in energie: hoe en met welke efficiëntie kunnen we één vorm van energie omzetten in een andere. We zeggen 'omzetten,' want energie verschijnt of verdwijnt niet zomaar: een kop koffie koelt af omdat ze energie afstaat aan de lucht van de kamer waarin ze staat; een motor zal energie die opgeslagen zit in brandstoffen omzetten naar bewegingsenergie, enzovoorts.

Daarom passen we statistiek toe op deze systemen (die uit enorm veel deeltjes bestaan), en we zijn dan bezig met 'statistische mechanica'.

Typisch gedrag en fluctuaties Ten eerste kunnen we statistiek gebruiken om te voorspellen wat zo'n systemen *typisch* doen, hun verwachte gedrag dus. Door dit te doen komen er begrippen tevoorschijn die we kennen in het dagelijkse leven, maar die niet echt bestaan of gewoon niet relevant zijn op het niveau van enkele atomen of moleculen. Temperatuur bijvoorbeeld: het blijkt dat temperatuur te maken heeft met hoeveel de deeltjes van een systeem gemiddeld bewegen. Verder blijkt dat er simpele relaties bestaan tussen begrippen als temperatuur, volume en energie. Zo vindt men bijvoorbeeld dat warme lucht meer energie bevat en een groter volume zal innemen (of een hogere druk zal hebben) dan koude lucht. De tak van de fysica die dit typische gedrag onderzoekt wordt thermodynamica genoemd. Historisch is ze ontstaan nog voor statistische mechanica, nog voor de wetenschap zelfs over atomen sprak. Ze ontstond in de negentiende eeuw, toen vooral efficiënties van motoren onderzocht werden.

Statistische mechanica biedt echter meer. Men kan hiermee ook bepalen wat voor afwijkingen een systeem vertoont op dit typisch gedrag. Zo'n afwijking wordt een *fluctuatie* genoemd. Het blijkt dat voor systemen met zoveel (10^{23}) deeltjes, deze fluctuaties doorgaans zo klein zijn dat we ze niet zien. Het is niet zo dat grotere fluctuaties onmogelijk zijn, ze zijn gewoon erg onwaarschijnlijk. Hoe groter het systeem, hoe minder fluctuaties we zien. Maar we moeten ons bedenken dat in de wereld van vandaag zeer kleine systemen ook van groot belang zijn. Denk maar aan nanotechnologie (computerchips enz) en biofysica (werking van cellen, enz). Zulke systemen hebben nog altijd veel deeltjes, zodat statistiek nodig is, maar zijn toch klein genoeg zodat fluctuaties van belang zijn. Verder zijn er ook chaotische systemen waar kleine afwijkingen enorme gevolgen kunnen hebben (het weer bijvoorbeeld), en er dus van typisch gedrag eigenlijk geen sprake is.

Uit evenwicht Zowel de 'klassieke' thermodynamica als de 'klassieke' statistische mechanica, daarmee bedoelen we de theorieën zoals die ontstaan zijn in de negentiende eeuw, zijn theorieën van systemen in evenwicht. Voor systemen uit evenwicht bestaat er vandaag geen alomvattende theorie, noch in thermodynamica noch in statistische mechanica.

Wat is het verschil tussen systemen die in evenwicht zijn en systemen die uit evenwicht zijn? Laat ons dit illustreren met een aantal voorbeelden:

-Neem een kop hete koffie en zet die op tafel. Zolang de koffie warmer is dan de lucht rondom, zal de koffie afkoelen. Ze zal dat doen door energie af te staan aan de lucht. De temperatuur van de koffie zal dus blijven dalen. Na verloop van tijd zal dit proces stoppen, dit is wanneer de koffie dezelfde temperatuur heeft als de kamer. In dat geval spreken we van evenwicht.

-Bekijk een muur van een huis, die aan de ene kant in contact staat met de koude buitenlucht en aan de andere kant met de verwarmde lucht in de kamer. Als die muur niet perfect isolerend is, dan zal er continu warmte van binnen naar buiten stromen. We spreken dan niet van evenwicht, en zolang het buiten koud blijft en

binnen de verwarming op volle toeren draait, zal er ook geen evenwicht zijn.
-Een heel ander voorbeeld: je zit in je kamer, terwijl het buiten waait. Zolang alle ramen en deuren dicht zijn, zal er niet veel gebeuren, er is evenwicht. Als je echter een raam en een deur openzet, kan er een luchtstroom door je kamer gaan en is er dus niet langer evenwicht.

Samengevat: zolang er stromen (uitwisselingen van energie) zijn tussen het systeem dat we bekijken en de rest van de wereld, zeggen we dat het systeem uit evenwicht is. Als er geen energiestromen zijn, is het systeem in evenwicht. Zoals al gezegd bestaat er alleen een alomvattende theorie voor systemen die in evenwicht zijn. Voor systemen uit evenwicht bestaat die nog niet. Dat een dergelijke theorie belangrijk is, is niet moeilijk om in te zien: zoals je je kunt bedenken zijn systemen in het dagelijks leven voortdurend uit evenwicht, zodat we een statistische mechanica nodig hebben die dergelijke processen verklaart vanuit een microscopisch startpunt. Maar ook fluctuaties zijn belangrijk, zeker voor kleine systemen. Een groot aantal van dit soort kleine systemen is ook inherent uit evenwicht. In de biofysica denken we bijvoorbeeld aan transport van deeltjes door celwanden en de ATP-cyclus die zorgt voor energie-opslag en -verbruik in ons lichaam.

Dit alles motiveert een zoektocht naar statistische mechanica voor systemen uit evenwicht. Deze thesis vat vier jaar van onderzoek samen, waarin we twee belangrijke deelgebieden van dit domein bestudeerd hebben: de respons van systemen op een perturbatie, samengevat in fluctuatie-dissipatie relaties, en dynamische fluctuaties, waarbij het onderzoek naar zogenaamde fluctuatiefunctionalen centraal staat.

D.2 Deel I: Stochastische processen

Om fysische systemen te beschrijven worden wiskundige modellen gebruikt. De modellen die wij gebruiken zijn stochastische processen. Ze beschrijven systemen op wat we noemen het mesoscopische niveau.

Mesoscopische systemen In de microscopische beschrijving van een systeem houden we rekening met alle posities en snelheden en alle krachten tussen alle deeltjes waaruit het systeem bestaat. Op het macroscopisch niveau zijn we alleen geïnteresseerd in een aantal variabelen als temperatuur en druk, en in het typische gedrag van het systeem (thermodynamica dus). In statistische mechanica werken we vaak op het mesoscopische niveau, als tussenstap tussen die twee. Op het mesoscopische beschrijvingsniveau beschrijft men systemen die van de ene kant veel deeltjes bevatten zodat we statistiek moeten toepassen, maar van de andere

kant ook niet zoveel deeltjes dat we ons alleen bezig moeten houden met het typische gedrag. Fluctuaties zijn dus heel belangrijk voor dergelijke systemen.

Een mooi voorbeeld is de waarneming die botanicus Robert Brown maakte in het begin van negentiende eeuw: hij bekeek door een microscoop hoe stuifmeelkorrels zich bewogen in een vloeistof. Doordat de stuifmeelkorrels voortdurend botsten met de moleculen van de vloeistof, maakten ze een zeer grillige beweging. Maar, met zijn microscoop kon hij niet de afzonderlijke moleculen van de vloeistof zien. Meer nog, hij wist zelfs niet dat vloeistof uit moleculen bestond. De stuifmeelkorrels maakten dus een zeer onvoorspelbare beweging. Onvoorspelbaar in de zin dat hij niet met zekerheid kon zeggen waarheen de korrels gingen bewegen. Over het typische gedrag van één korrel viel dus niet veel te zeggen, het waren de fluctuaties die interessant waren.

In het eerste deel van de thesis gaan we in op de modellen die we gebruikt hebben om fysische mesoscopische systemen te beschrijven: stochastische processen. De belangrijke fysische ingrediënten in deze beschrijving zijn dissipatie en activiteit.

Dissipatie en activiteit Systemen die niet in evenwicht zijn, wisselen voortdurend energie uit met hun omgeving. Eén vorm van energieuitwisseling is warmte. Warmtestromen krijg je als je twee voorwerpen die een verschillende temperatuur hebben met elkaar in contact brengt: de energie die dan van warm naar koud stroomt noemen we warmte. Als een systeem warmte uitwisselt met zijn omgeving dan spreken we van energiedissipatie. In het dagelijkse leven kan dat een voordeel zijn (voor een verwarming bijvoorbeeld) of een nadeel (voor een gloeilamp beschouwen we warmte als energieverlies). Een fysische grootheid die daar veel mee te maken heeft is entropie. Dat is een centraal begrip in statistische mechanica en thermodynamica. Het is ook één van de hoofdingrediënten in de modellen die we gebruiken. Het voordeel van entropie/dissipatie is dat dit in experimenten daadwerkelijk gemeten kan worden.

Het andere hoofdingrediënt is de activiteit van het systeem. Met de activiteit bedoelen we ongeveer het volgende: het geeft weer hoeveel of hoe vaak het systeem verandert in de tijd. In het voorbeeld van de stuifmeelkorrel spreken we van een grote activiteit als de korrel heel veel beweegt, hoe 'wilder' hoe groter de activiteit. Het probleem met deze activiteit is dat het moeilijk meetbaar is op het thermodynamisch niveau. Tenminste tot nu toe.

D.3 Deel II: Fluctuatie-dissipatie relaties

Hoe reageren systemen op een perturbatie? Een perturbatie is een kleine verandering die we opleggen aan het systeem. In ons geval voegen we een klein

beetje energie toe.

Laat ons een voorbeeld nemen om dit te verduidelijken. Neem terug de stuifmeelkorrels die in een vloeistof bewegen. Een beetje energie toevoegen komt hier overeen met een kleine kracht uitoefenen op zo'n korrel. Stel je bijvoorbeeld voor dat we een heel dun draadje aan de korrel vastmaken en daaraan trekken. We zijn dan bijvoorbeeld geïnteresseerd in de gemiddelde snelheid die de korrel krijgt. Hoe groter de wrijving is die de korrel ondervindt van de vloeistof, hoe kleiner de snelheid. Maar wrijving is het gevolg van botsingen van de korrel met de vloeistofmoleculen. Van de andere kant zijn die botsingen (ook als we geen kracht uitoefenen) ook verantwoordelijk voor het onvoorspelbare gedrag van de korrel, dus van de fluctuaties.

Dit blijkt een algemeen gegeven te zijn: de mate waarin het systeem reageert op een perturbatie, hangt nauw samen met de fluctuaties die het systeem vertoont. In het algemeen bekijken we de typische waarden (verwachtingswaarden) van meetbare grootheden (observabelen). De mate waarin die verwachtingswaarden veranderen wordt de respons genoemd. De fluctuaties die het systeem vertoont kunnen in zekere mate berekend worden door correlatiefuncties te bekijken. Een correlatiefunctie is een maat voor het verband tussen de waarden van twee observabelen.

Voor systemen in evenwicht weet men dat de respons kan uitgedrukt worden in termen van een correlatiefunctie tussen de observabele en de energie die de perturbatie vormt. Dit verband wordt het fluctuatie-dissipatie theorema genoemd.

Onze contributie is het onderzoeken van de respons voor systemen uit evenwicht. We bespreken in dit deel van de thesis een algemene strategie om respons uit te rekenen. Deze strategie is eerder gepubliceerd in [3]. We zien dat de respons uitgedrukt kan worden als de som van twee correlatiefuncties. Eentje is de correlatie van de observabele met entropie (dus energiedissipatie), en de andere met de activiteit van het systeem. We leiden ook meer concrete resultaten af voor een aantal klassen van modellen, zie ook [2, 4].

D.4 Deel III: Dynamische fluctuaties

In dit deel van de thesis bekijken we de kans op grote afwijkingen van het typisch gedrag van systemen. Hiervoor bestaat er in de wiskunde een goedbestudeerde theorie, namelijk die van de grote afwijkingen. Wanneer deze toegepast wordt op de statistische mechanica van systemen in evenwicht, blijkt dat ze een natuurlijk wiskundig kader vormt om de fysische theorie in te beschrijven. Het is dan ook logisch om deze theorie te gebruiken om systemen uit evenwicht te bestuderen.

In [70, 71] bestuderen we deze grote afwijkingen voor verschillende klassen van modellen. De grote lijnen zijn in beide gevallen echter dezelfde. Ten eerste beschouwen we twee soorten observabelen: occupaties (die beschrijven hoe lang het systeem in elke toestand verblijft) en stromen (die beschrijven hoeveel het systeem van toestand verandert).

We bestuderen dan de kans dat deze observabelen waarden aannemen verschillend van de typische (verwachte) waarden. Dergelijke kansen worden vaak exponentieel kleiner naarmate de tijd vordert. Een functie die een maat is voor hoe snel dit 'exponentieel verval' gebeurt is de fluctuatiefunctionaal, en die staat centraal in dit deel van de thesis.

De vragen die we ons hier stellen zijn: hoe kunnen we die fluctuatiefunctionaal berekenen en wat zijn de fysische grootheden die bepalend zijn voor zijn vorm? Net zoals in deel 2 zijn entropie en activiteit hier de belangrijke ingrediënten. De fluctuatiefunctionaal kan expliciet berekend worden voor beide klassen van modellen die we beschouwen, en kan uitgedrukt worden in termen van entropie en activiteit.

Wanneer we de fluctuatiefunctionaal berekenen voor de occupaties alleen, zien we dat die volledig uitgedrukt kan worden in termen van de activiteit. We leggen ook een structuur bloot die sterk doet denken aan de relaties tussen grootheden uit de thermodynamica (thermodynamische potentialen). Tenslotte, voor kleine fluctuaties kunnen we een aantal gekende fysische principes terugvinden, die te maken hebben met entropie.

We besluiten hier dat de activiteit essentieel is in niet-evenwichts statistische mechanica. Vooral dynamische fluctuaties geven een beter inzicht in wat deze grootheid kan zijn: enerzijds een maat voor de ontsnappingskans van een systeem uit zijn toestand, anderzijds als activiteit, het aantal keren dat het systeem van toestand verandert.

Bibliography

[1] G. S. Agarwal. Fluctuation-dissipation theorems for systems in non-thermal equilibrium and applications. *Z. Physik*, 252:25–38, 1972.

[2] M. Baiesi, E. Boksenbojm, C. Maes, and B. Wynants. Nonequilibrium linear response for Markov dynamics, II: Inertial dynamics. *J. Stat. Phys.*, 139:492–505, 2010.

[3] M. Baiesi, C. Maes, and B. Wynants. Fluctuations and response of nonequilibrium states. *Phys. Rev. Lett.*, 103:010602, 2009.

[4] M. Baiesi, C. Maes, and B. Wynants. Nonequilibrium linear response for Markov dynamics, I: Jump processes and overdamped diffusions. *J. Stat. Phys.*, 137(5-6):1094–1116, 2009.

[5] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Fluctuations in stationary nonequilibrium states of irreversible processes. *Phys. Rev. Lett.*, 87:040601, 2001.

[6] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Macroscopic fluctuation theory for stationary non equilibrium states. *J. Stat. Phys.*, 107:635–675, 2002.

[7] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Current fluctuations in stochastic lattice gases. *Phys. Rev. Lett.*, 94:030601, 2005.

[8] W. Blickle, T. Speck, C. Lutz, U. Seifert, and C. Bechinger. Einstein relation generalized to nonequilibrium. *Phys. Rev. Lett.*, 98:210601, 2007.

[9] T. Bodineau and B. Derrida. Current fluctuations in non-equilibrium diffusive systems: an additivity principle. *Phys. Rev. Lett.*, 92:180601, 2004.

[10] T. Bodineau and R. Lefevere. Large deviations of lattice hamiltonian dynamics coupled to stochastic thermostats. *J. Stat. Phys.*, 133(1):1–27, 2008.

[11] S. Bruers, C. Maes, and K. Netočný. On the validity of entropy production principles for linear electrical circuits. *J. Stat. Phys.*, 129(4):725–740, 2007.

[12] P. Calabrese and A. Gambassi. Ageing properties of critical systems. *J. Phys. A: Math. Gen.*, 38:R133–R193, 2005.

[13] H. B. Callen and T. A. Welton. Irreversibility and generalized noise. *Phys. Rev.*, 83:34–40, 1951.

[14] C. Chatelain. A far-from-equilibrium fluctuation-dissipation relation for an Ising-Glauber-like model. *J. Phys. A*, 36:10739, 2003.

[15] R. Chetrite, G. Falkovich, and K. Gawędzki. Fluctuation relations in simple examples of non-equilibrium steady states. *J. Stat. Mech.*, page P08005, 2008.

[16] R. Chetrite and K. Gawędzki. Eulerian and lagrangian pictures of non-equilibrium diffusions. *J. Stat. Phys.*, 137:890–916, 2009.

[17] A. Chrisanti and J. Ritort. Violation of the fluctuation-dissipation theorem in glassy systems: basic notions and the numerical evidence. *J. Phys. A: Math. Gen.*, 36:R181–R290, 2003.

[18] W. T. Coffey, Y. P. Kalmykov, and J. T. Waldron. *The Langevin Equation; With Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering*. World Scientific Publishing Co., 2nd edition, 2004.

[19] G. E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. *J. Stat. Phys.*, 90:1481–1487, 1998.

[20] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E*, 60:2721–2726, 1999.

[21] L. Cugliandolo, J. Kurchan, and G. Parisi. Off equilibrium dynamics and aging in unfrustrated systems. *J. Phys. I*, 4:1641, 1994.

[22] W. De Roeck and C. Maes. Symmetries of the ratchet current. *Phys. Rev. E*, 76:051117, 2007.

[23] U. Deker and F. Haake. Fluctuation-dissipation theorems for classical processes. *Phys. Rev. A*, 11:2043, 1975.

[24] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer-Verlag, New York, 2nd edition, 1998.

[25] B. Derrida. Non equilibrium steady states: fluctuations and large deviations of the density and of the current. *J. Stat. Mech.*, page P07023, 2007.

[26] G. Diezemann. Fluctuation-dissipation relations for Markov jump processes. *Phys. Rev. E*, 72:011104, 2005.

[27] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. *Comm. Pure Appl. Math.*, 28:1–47, 1975.

[28] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. II. *Comm. Pure Appl. Math.*, 28:279–301, 1975.

[29] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. III. *Comm. Pure Appl. Math.*, 29:389–461, 1976.

[30] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. IV. *Comm. Pure Appl. Math.*, 36:183–212, 1983.

[31] J. R. Dorfman. *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*. Cambridge University Press, 1999.

[32] J. R. Dorfman and P. Gaspard. Chaotic scattering theory of transport and reaction-rate coefficients. *Phys. Rev. E*, 51:28–35, 1995.

[33] A. Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik:*, 17, 549–560.

[34] R. S. Ellis. Large deviations for a general class of random vectors. *Ann. Probab.*, 12(1):1–12, 1984.

[35] R. S. Ellis. *Entropy, Large Deviations and Statistical Mechanics*. Springer, New-York, 1985.

[36] R. S. Ellis. An overview of the theory of large deviations and applications to statistical mechanics. *Scand. Actuar. J.*, 1:97–142, 1995.

[37] R. S. Ellis. The theory of large deviations: from BoltzmannâĂŹs 1877 calculation to equilibrium macrostates in 2D turbulence. *Physica D*, 133:106–136, 1999.

[38] D.J. Evans, E. G. D. Cohen, and G. P. Morris. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.*, 71:2401–2404, 1993.

[39] D.J. Evans and D. J. Searles. Equilibrium microstates which generate second law violating steady states. *Phys. Rev. E*, 50:1645–1648, 1994.

[40] M. Falcioni, S. Isola, and A. Vulpiani. Correlation functions and relaxation properties in chaotic dynamics and statistical mechanics. *Phys. Lett. A*, 144:341, 1990.

[41] M.I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York, 1984.

[42] G. Gallavotti and E. G. D. Cohen. Dynamical ensembles in nonequilibrium statistical mechanics. *J. Stat. Phys.*, 80:2694–2697, 1995.

[43] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and The Natural Sciences*. Springer, Berlin, 2004.

[44] J. Gärtner. On large deviations from the invariant measure. *Theory Probab. Appl.*, 22:24–39, 1977.

[45] J. R. Gomez-Solano, A. Petrosyan, S. Ciliberto, R. Chetrite, and K. Gawędzki. Experimental verification of a modified fluctuation-dissipation relation for a micron-sized particle in a non-equilibrium steady state. *Phys. Rev. Lett.*, 103:075007, 2009.

[46] G. Grimmett and D. Stirzaker. *Probability and Random processes*. Oxford University Press, USA, 3rd edition, 2001.

[47] P. Hänggi and H. Thomas. Linear response and fluctuation theorems for nonstationary stochastic processes. *Z. Physik B*, 22:295–300, 1975.

[48] P. Hänggi and H. Thomas. Stochastic processes: time-evolution, symmetries and linear response. *Phys. Rep.*, 88:207–319, 1982.

[49] T. Harada and S. Y. Sasa. Equality connecting energy dissipation with violation of fluctuation-response relation. *Phys. Rev. Lett.*, 95:130602, 2005.

[50] T. Hatano and S. Y. Sasa. Steady-state thermodynamics of Langevin systems. *Phys. Rev. Lett.*, 86:3463, 2001.

[51] M. Henkel and M. Pleimling. *Non-Equilibrium Phase Transitions: Volume 2: Dynamical Scaling far from Equilibrium*. Canopus Academic Publishing Limited, 2009.

[52] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997.

[53] C. Jarzynski. Microscopic analysis of Clausius-Duhem processes. *J. Stat. Phys.*, 96(1-2):415–427, 1999.

[54] C. Jarzynski. Hamiltonian derivation of a detailed fluctuation theorem. *J. Stat. Phys.*, 98(1-2):77–102, 2000.

[55] J. Johnson. Thermal agitation of electricity in conductors. *Phys. Rev.*, 32:97–109, 1928.

[56] S. Katz, J. L. Lebowitz, and H. Spohn. Phase transitions in stationary nonequilibrium states of model lattice systems. *Phys. Rev. B*, 28(3):1655–1658, 1983.

[57] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems*. Springer-Verlag, Berlin, 1999.

[58] N. V. Krylov. *Introduction to the Theory of Diffusion Processes*. American Mathematical Society, 1995.

[59] R. Kubo. The fluctuation-dissipation theorem. *Rep. Prog. Phys.*, 29:255–284, 1966.

[60] J. Kurchan. In and out of equilibrium. *Nature*, 433:222, 2005.

[61] O. E. Lanford. *Entropy and Equilibrium States in Classical Statistical Mechanics*, pages 1–113. Springer, Berlin, 1973. in: Statistical Mechanics and Mathematical Problems, in: Lecture notes in Physics.

[62] J. L. Lebowitz. Stationary nonequilibrium Gibbsian ensembles. *Phys. Rev.*, 114(5):1192–1202, 1959.

[63] E. Lippiello, F. Corberi, A. Saraccino, and M. Zanetti. Nonlinear response and fluctuation dissipation relations. *Phys. Rev. E*, 78:041120, 2008.

[64] E. Lippiello, F. Corberi, and M. Zanetti. Off-equilibrium generalization of the fluctuation dissipation theorem for ising spins and measurement of the linear response. *Phys. Rev. E*, 71:036104, 2005.

[65] R. S. Lipster and A. N. Shiryayev. *Statistics of Random Processes*, volume I, II. Springer-Verlag, New York, 1978.

[66] C. Maes. Fluctuation theorem as a Gibbs property. *J. Stat. Phys.*, 95:367–392, 1999.

[67] C. Maes and K. Netočný. Time reversal and entropy. *J. Stat. Phys.*, 110:269–310, 2003.

[68] C. Maes and K. Netočný. Minimum entropy production principle from a dynamical fluctuation law. *J. Math. Phys.*, 48:053306, 2007.

[69] C. Maes and K. Netočný. Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. *Europhys. Lett.*, 82:30003, 2008.

[70] C. Maes, K. Netočný, and B. Wynants. On and beyond entropy production: the case of Markov jump processes. *Markov Processes Relat. Fields*, 14:445–464, 2008.

[71] C. Maes, K. Netočný, and B. Wynants. Steady state statistics of driven diffusions. *Physica A*, 387:2675–2689, 2008.

[72] C. Maes, K. Netočný, and B. Wynants. Dynamical fluctuations for semi-Markov processes. *J. Phys. A: Math. Theor.*, 42:365002, 2009.

[73] C. Maes, F. Redig, and A. Van Moffaert. On the definition of entropy production via examples. *J. Math. Phys.*, 41:1528–1554, 2000.

[74] C. Maes, F. Redig, and M. Verschuere. From global to local fluctuation theorems. *Moscow Mathematical Journal*, 1:421–438, 2001.

[75] C. Maes and B. Wynants. On a response formula and its interpretation. *arXiv:0910.2320*, 2009.

[76] U. Marini Bettolo Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani. Fluctuation–dissipation: Response theory in statistical physics. *Phys. Rep.*, 461(4-6):111–195, 2008.

[77] K. Martens, E. Bertin, and M. Droz. Dependence of the fluctuation-dissipation temperature on the choice of observable. *Phys. Rev. Lett.*, 103:260602, 2009.

[78] P. A. Martin. Physique statistique des processus irreversibles. Lecture notes of the DEA de Physique Theorique, notes by F. Coppex, 2004.

[79] J. A. McLennan Jr. Statistical mechanics of the steady state. *Phys. Rev.*, 115:1405–1409, 1959.

[80] J. A. McLennan Jr. *Introduction to Nonequilibrium Statistical Mechanics*. Prentice-Hall, En- glewood Cliffs, NJ, 1989.

[81] R. E. Mortensen. Mathematical problems of modeling stochastic nonlinear dynamic systems. *J. Stat. Phys.*, 1(2):271–296, 1969.

[82] T. Nakamura and S. Y. Sasa. A fluctuation-response relation of many Brownian particles under nonequilibrium conditions. *Phys. Rev. E*, 77:021108, 2008.

[83] O. Nikodym. Sur une généralisation des intégrales de M. J. Radon. *Fundamenta Mathematicae*, 15:131–179, 1930.

[84] H. Nyquist. Thermal agitation of electric charge in conductors. *Phys. Rev.*, 32:110–113, 1928.

[85] L. Onsager. Reciprocal relations in irreversible processes. I. *Phys. Rev.*, 37:405–426, 1931.

[86] L. Onsager. Reciprocal relations in irreversible processes. II. *Phys. Rev.*, 38:2265–2279, 1931.

[87] L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Phys. Rev.*, 91(6):1505–1512, 1953.

[88] Y. Oono and M. Paniconi. Steady state thermodynamics. *Prog. Theor. Phys. Suppl.*, 130:29, 1998.

[89] I. Prigogine. *Introduction to Non-Equilibrium Thermodynamics.* Wiley-Interscience, New York, 1962.

[90] A. Rákos and R. J. Harris. On the range of validity of the fluctuation theorem for stochastic markovian dynamics. *J. Stat. Mech.*, page P05005, 2008.

[91] F. Ricci-Tersenghi. Measuring the fluctuation-dissipation ratio in glassy systems with no perturbing field. *Phys. Rev. E*, 68:065104(R), 2003.

[92] F. Ritort. Universal dependence of the fluctuation-dissipation ratio on the transition rates in trap models. *J. Phys. A: Math. Gen.*, 36:10791–10805, 2003.

[93] D. Ruelle. A review of linear response theory for general differential dynamical systems. *Nonlinearity*, 22:855–870, 2009.

[94] J. M. Sancho, M. San Miguel, and D. Dürr. Adiabatic elimination for systems of Brownian particles with nonconstant damping coefficients. *J. Stat. Phys.*, 28(2):291–305, 1982.

[95] J. P. Sethna. *Entropy, Order Parameters, and Complexity.* Oxford Univ. Press, Oxford, 2006.

[96] N. Singh and B. Wynants. Dynamical fluctuations for periodically driven diffusions. *J. Stat. Mech.*, page P03007, 2010.

[97] T. Speck and U. Seifert. Restoring a fluctuation-dissipation theorem in a nonequilibrium steady state. *Europhys. Lett.*, 74:391–396, 2006.

[98] T. Speck and U. Seifert. Extended fluctuation-dissipation theorem for soft matter in stationary flow. *Phys. Rev. E*, 79:040102, 2009.

[99] U. M. Titulaer. A systematic solution procedure for the Fokker-Planck equation of a brownian particle in the high-friction case. *Physica*, 91A:321–344, 1978.

[100] H. Touchette. The large deviation approach to statistical mechanics. *Phys. Rep.*, 478:1–69, 2009.

[101] G. Volpe, L. Helden, T. Brettschneider, J. Wehr, and C. Bechinger. Influence of noise on force measurements. *arXiv:1004.0874v1*, 2010.

[102] W. Weidlich. Fluctuation-dissipation theorem for a class of stationary open systems. *Z. Physik*, 248:234–243, 1971.

[103] A. D. Wentzell and M.I. Freidlin. On small random perturbations of dynamical systems. *Russ. Math. Surveys*, 25:1–55, 1970.

[104] G. Wilemski. In the derivation of Smoluchowski equations with corrections in the classical theory of Brownian motion. *J. Stat. Phys.*, 14(2):153–169, 1976.

Curriculum vitae

Personal data:

Date of birth	31 - 03 - 1984
Nationality	Belgian
E-mail	bram.wynants@fys.kuleuven.be
Address	Celestijnenlaan 200D - 06.10
	3001 Heverlee
	Belgium
Tel.	+32473861031

Education:

Undergraduate studies	Physics at the K.U.Leuven, 2002-2006, Leuven, Belgium
PhD	Theoretical physics at the K.U.Leuven, 2006-2010, Leuven, Belgium
Fellowship	aspirant of FWO flanders (www.fwo.be)

Teaching experience:

Teacher assistant for several courses on general physics for students of medical sciences, under Prof. Jos Rogiers, and thermophysics for physics students under Prof. Christian Maes.

Seminars given at:

Academy of Sciences, Prague; Max-Planck Institute, Dresden; K.U.Leuven, Leuven, Universiteit Antwerpen, Antwerpen; Palais des Académies, Brussels; U.C.Louvain, Louvain-La-Neuve.

Published and submitted articles (Authors in alphabetical order):

1. E. Boksenbojm, B. Wynants, C. Jarzynski,
   Nonequilibrium thermodynamics at the microscale: Work relations and the second law,
   lecture notes for lectures given by C.J. at Fundamental Problems in Statistical Physics XII, Leuven, Aug 30 - Sept 12, 2009, accepted for publication in Physica A, arXiv:1002.1230

2. M. Baiesi, E. Boksenbojm, C. Maes, B. Wynants,
   Nonequilibrium Linear Response for Markov Dynamics, II: Inertial Dynamics,
   Journal of statistical physics **139**, 492-505 (2010)

3. N. Singh, B. Wynants
   Dynamical fluctuations for periodically driven diffusions,
   Journal of Statistical Mechanics (2010) P03007

4. C. Maes, B. Wynants
   On a response formula and its interpretation,
   Markov Processes and Related Fields **16**, 45-58 (2010)

5. M. Baiesi, C. Maes, B. Wynants:
   Nonequilibrium linear response for Markov dynamics, I: jump processes and overdamped diffusions,
   Journal of Statistical Physics **137**, 1094-1116 (2009)

6. E. Boksenbojm, B. Wynants:
   Entropy and efficiency of a molecular motor model,
   Journal of Physics A: Mathematical and Theoretical **42** (2009) 445003

7. K. Netočný, C. Maes, B. Wynants:
   Dynamical fluctuations for semi-Markov processes,
   Journal of Physics A: Mathematical and Theoretical **42**, (2009) 365002

8. M. Baiesi, C. Maes, B. Wynants:
   Fluctuations and response of nonequilibrium states,
   Physical Review Letters **103**, 010602 (2009)

9. K. Netočný, C. Maes, B. Wynants:
   On and beyond entropy production: the case of Markov jump processes,
   Markov Processes and Related Fields **14**, 445–464 (2008)

10. K. Netočný, C. Maes, B. Wynants:
    Steady state statistics of driven diffusions,
    Physica A **387**, 2675–2689 (2008)

Index